

Tailoring conservative care in osteoarthritis

Citation for published version (APA):

Hoogeboom, T. J. (2012). *Tailoring conservative care in osteoarthritis*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20121130th>

Document status and date:

Published: 01/01/2012

DOI:

[10.26481/dis.20121130th](https://doi.org/10.26481/dis.20121130th)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Tailoring conservative care in osteoarthritis

ISBN: 978-94-6182-179-9

Printing: Off Page, www.offpage.nl

Cover concept: A. Sauvage, www.asauvage.com

Cover design: Sanne Smeekeens

Photograph: Dylan Don, www.dylandon.com

The publication of this thesis was financially supported by (in alphabetical order) the Annafonds|NOREF (Leiden), Dutch Arthritis Foundation (Amsterdam), and the Scientific College of Physiotherapy (WCF) (Amersfoort).

Copyright © 2011 by Thomas Hoogeboom. All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without permission of the author.

Tailoring conservative care in osteoarthritis

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. mr. G.P.M.F. Mols,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 30 november 2012 om 12.00 uur

door

Thomas Johannes Hoogeboom

Promotor:

Prof. dr. R.A. de Bie

Copromotores:

Dr. C.H.M. van den Ende

Dr. A.A. den Broeder

Beoordelingscommissie:

Prof. dr. S. van der Linden (Voorzitter)

Prof. dr. S.M.A. Bierma-Zeinstra (Erasmus MC)

Prof. dr. J.A. Knottnerus

Dr. N.L.U. van Meeteren (TNO)

Prof. dr. M.H. Prins

CONTENTS

Chapter 1:	General introduction	7
	<i>Part I: Assessing the needs of people with osteoarthritis</i>	
Chapter 2:	The highly reliable, valid and responsive Dutch lower extremity functional scale is the preferred measure for functioning in hip or knee osteoarthritis: a validation study	21
Chapter 3:	Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study	39
Chapter 4:	Longitudinal study of the impact of joint-pain comorbidities on quality of life and activity levels in people with knee osteoarthritis: Data from the Osteoarthritis Initiative	49
Chapter 5:	The impact of waiting for total joint replacement on pain and functional status: a systematic review	59
Chapter 6:	Prevalence and predictors of health care use in patients with early hip or knee osteoarthritis: two-year follow-up data from the CHECK cohort	81
	<i>Part II: Tailoring care for people with advanced osteoarthritis</i>	
Chapter 7	A multi-disciplinary and multidimensional intervention for patients with hand osteoarthritis	99
Chapter 8:	Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial	117
Chapter 9:	Piloting and preliminary evaluating a multi-disciplinary care program for individuals with generalized osteoarthritis on pain and self-efficacy: a concurrent multiple-baseline single case study	139
Chapter 10:	Effectiveness of preoperative therapeutic exercise for major joint replacement on postoperative functional recovery: a meta-analysis and best practice evaluation	155
Chapter 11:	Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized clinical pilot trial	183
Chapter 12:	General discussion	199
Addendum:	Summary	213
	Samenvatting	217
	Dankwoord	221
	About the author	225
	List of publications	229

This page intentionally left blank

1

General introduction

INTRODUCTION

The purpose of this thesis is to assess the *health needs* of people with osteoarthritis and to evaluate non-pharmacological interventions *tailored* to those needs. This introduction chapter outlines the epidemiology, pathology and clinical features of osteoarthritis. Furthermore, it details on the current management of people with osteoarthritis and discusses its potential pitfalls and opportunities. Finally, it introduces the research questions this thesis is aiming to answer.

Current knowledge on osteoarthritis

Epidemiology

Osteoarthritis is a prevalent medical condition, with an estimated number of affected patients in the Netherlands of 240.000 men and 417.000 women (1). In general, four subgroups of osteoarthritis have been established: 1) knee, 2) hip, 3) hand, and 4) generalized osteoarthritis¹ (GOA). Overall prevalence of osteoarthritis (clinical and radiological) (95%-CI) of the knee is estimated to be 23.9% (23.6 – 24.2), for the hip 10.9% (10.6-11.2) and for the hand 43.3% (42.7-42.9), in people over 40 years of age (2). The prevalence of GOA (depending on definition) ranges from 27% (3) to 52% (4) in a sample of individuals with hip or knee osteoarthritis. Due to demographic changes and the increasing prevalence of obesity, the prevalence of osteoarthritis is expected to grow in the next three decades (5). Osteoarthritis generally develops progressively over several years, although symptoms might remain stable for long periods or even improve (6). Diagnosing osteoarthritis relies on the assessment of clinical and radiological features (7). Interestingly, nearly half of patients with radiological features of osteoarthritis experience no symptoms and vice versa (8;9). Risk factors for primary osteoarthritis include old age, high bone mass, a genetic predisposition for the disease, increased body mass index, participation in weight-bearing sports (e.g., running at elite level), and occupations that require prolonged standing, lifting, or moving of heavy objects (10-12).

Pathology

Osteoarthritis has mainly been portrayed as a failure of the repair process of damaged cartilage due to biomechanical changes in the joint. Cartilage is a non-vascularised tissue and is therefore dependent of nutrient and oxygen supply by chondrocytes. Chondrocytes are responsible for the maintenance of a very large proportion of the extracellular matrix. In the earliest stage of cartilage damage, clusters of chondrocytes form in the damaged areas

¹ Many different definitions for GOA have been proposed. In this thesis GOA is defined as having complaints in three different joint groups and osteoarthritis signs in at least two joints.

and the concentration of growth factors in the matrix rises in a repair attempt (13;14). Subsequently, this attempt fails and results in an imbalance in favour of degradation (15;16), eventually resulting in the formation of a matrix that is unable to withstand normal mechanical stresses. Consequently, the tissue enters a vicious cycle in which breakdown dominates synthesis of the extracellular matrix (17). However, the loss of cartilage – by its attempts to regenerate itself – is also accompanied by joint remodeling, for instance by new bone formation (osteophytes). In addition, joint inflammation occurs which is influenced by local structures including ligaments, bones, tendons and effusions (18).

As articular cartilage is aneural, these changes do not produce clinical signs unless innervated tissue becomes involved. Recent research advances demonstrate that synovitis and bone marrow lesions are important factors contributing to pain in osteoarthritis (19;20). Inflammation of the synovial tissue (i.e. synovitis) corresponds to clinical symptoms such as joint swelling and inflammatory pain. Bone marrow lesions may appear after acute injury (21) or after chronic excess loading (22) and are more often found in individuals with painful knee osteoarthritis than in individuals with non-painful osteoarthritis, thus suggesting a causal role in the development of pain (23;24). Neuroimaging has provided evidence of the contribution of the central brain network to pain perception in osteoarthritis. In a study by Baliki *et al* (2008) local treatment of the osteoarthritic knee resulted in reduction of brain activity, suggesting that central activation of the brain mediates pain during osteoarthritis (25). Moreover, data by Kulkarni *et al* propose that arthritic pain is associated with areas of the brain responsible for affect, aversive conditioning and motivation (26). These findings may have important implications for the management of patients with osteoarthritis pain (20).

Clinical features

Pain is the most important symptom of osteoarthritis. People with hip and knee osteoarthritis experience two distinct types of pain: a dull aching pain that is more consistent over time; and, intermittent pain that is often more intense, unpredictable and emotionally draining (27). The pain experienced is typically worst during and after weight-bearing activities. Patients with osteoarthritis also experience stiffness; generally resolving within 30 minutes. Inflammatory flares occur during the course of the disease. Another important symptom of osteoarthritis is the loss of function and limitations in daily activities, such as walking, doing household chores and stair climbing (28). Factors associated with future decline in functional status comprise avoidance of activity, increased pain, increased stiffness, more comorbidity, higher age, longer disease duration, reduced muscle strength, reduced knee stability and range of joint motion (10;28-30). Other symptoms often associated with osteoarthritis are depression, fatigue and disturbed sleep; all contributing to the disability of osteoarthritis and diminishing of patients' quality of life.

Current evidence-based practice

To date, several guidelines are available for the management of hip and knee osteoarthritis (31-33). The general tendency in each of these guidelines is that joint replacement surgery should be postponed as long as possible, as the median life spans of these prostheses are approximately 15 years (31) and revision surgery is not often without risks or disappointing outcome. Postponing joint replacement surgery can be done with a whole spectrum of non-surgical therapy modalities. In 2011, Smink *et al* published a step-wise care strategy – based on the latest guideline recommendations for non-surgical treatment options – to guide health professionals and patients to manage the disease in an efficient, equitable and thorough fashion (34). Their stepped-care strategy implies that the more advanced options should only be considered if prior options failed, although the validity of this assumption remains to be tested.

Table 1. Step-wise representation of recommendations for the therapy management of osteoarthritis, adapted from the study of Smink *et al* (2011).

	Step 1	Step 2	Step 3
Treatment modalities	<ul style="list-style-type: none"> • Education • Lifestyle advice • Medication* <ul style="list-style-type: none"> - Acetaminophen 	<ul style="list-style-type: none"> • Exercise therapy • Dietary therapy • Medication* <ul style="list-style-type: none"> - (Topical) NSAIDs - Tramadol 	<ul style="list-style-type: none"> • Multidisciplinary care • TENS • Medication* <ul style="list-style-type: none"> - Intra-articular injections

* Consult current guidelines for an adequate dose.

In the BART-model therapy, recommendations are described in three different steps. In the first stage of the disease patients generally visit their primary care physician for (mild) localized discomfort of the hip or knee. Treatment options in the first step comprise education, lifestyle advice and adequate prescription of acetaminophen. In the second step, patients are likely to experience complaints for a considerable time and patients expect more intensive therapy measures. Treatment modalities that should be considered are exercise therapy, NSAIDs, and tramadol. Obese patients should be motivated to lose weight. If the treatment modalities in the previous steps do not elicit any beneficial effects, more advanced treatment options should be considered, i.e. multidisciplinary care, intra-articular injections with hyaluronic acid or glucocorticoids, and TENS. Finally, patients, who do not experience symptom-relieve from adequate non-surgical management, should be considered for joint replacement surgery (31;33;35).

What this thesis adds

As described in the previous paragraph, multiple therapy options are available in the management of osteoarthritis. As usual this is a sign that no single conservative treatment option is the ideal solution for all individuals with osteoarthritis in terms of efficacy, safety, or feasibility. Indeed, effects of these therapies are mostly small to modest (31;33;36-39). Non-pharmacological treatment options such as exercise, joint protection, weight reduction, and other measures to unload damaged joints are generally thought to elicit only small effects (effect size 0.20–0.50) (31;32) or in some cases no effects at all (39). Also, joint replacement surgery outcomes can be disappointing, as a significant number of patients experience continuing pain and functional disability after surgery (40;41). As yet, when therapy results are disappointing, health professionals and patients often implement several interventions or combinations of interventions to achieve symptomatic relief (42). This could result in an increased chance of adverse events and therapy costs (42;43). On the other hand, data from sub-group analyses of randomized clinical trials demonstrate that some patients respond differently to therapy interventions than others (44-46), which indicates that patient characteristics (such as, age, duration of symptoms, pain levels, physical functioning, and Kellgren and Lawrence grade) play an important role in the effectiveness of therapies. These findings are in line with the general belief that the variation in therapy results is most likely explained by the (clinical) heterogeneity of osteoarthritis (20;47-49). So, even though on group level the therapy effects are small to modest, on individual level benefits from therapy approaches may differ because of differences in *health needs*. The value of more personalized therapy options should therefore be explored in the management of osteoarthritis. To do so, researchers and health professionals should determine which aspects of the disease contribute to the (clinical) heterogeneity of osteoarthritis, through *health needs* assessments (Part 1 of this thesis). Consequently, health professionals should individualize their therapy options accordingly, through *tailored care* (Part 2 of this thesis) (50).

Health needs

When studying *health needs*, one should be aware that this comprises many different aspects of health. Health needs can be defined as: 'the patient's ability to benefit from health care services' (51). This definition might imply a focus on 'health' rather than 'health care', but one should also take factors such as socio-economic status and the variation in knowledge and attitudes of the population into account. Ideally, the provision of health care services meets most of the populations' needs. So when important health (care) needs are identified, one can tailor the therapeutic intervention to these needs. So through tailored care one might be able to maximise the therapeutic intervention's effectiveness by systematic analysis of the person's health and demographic and clinical profile.

Tailored care

Tailoring care can be done on two different levels:

1. Group level; by setting strict selection criteria for the intervention that ensure that the characteristics and capacities of the participants match the aims of the therapy modalities; resulting in a subgroup of patients that is most likely to benefit from treatment.
2. Patient level; by setting personal goals for each of the participants to work on. Even if the patient characteristics are matched on group level, it is unlikely that on an individual level personal goals match between patients. Therefore, interventions should be personalized and contextualized to the individual. This can be done through methods such as goal setting [65], which allows participants to set goals on a personal level and then work towards them.

In sum, the current conservative management of osteoarthritis might improve its effectiveness when clinically relevant health needs (on group and patient level) of individuals with osteoarthritis are taken into account, by tailoring care towards those needs.

Content of thesis

To enable tailored care, thorough assessment of the patient's *needs* and capacities preceding a therapy program is necessary. Therefore, in the first part of this thesis the aim was to identify *health needs* that have a clinically relevant influence on the health outcome in the course of osteoarthritis. In the second part of the thesis the aim was to evaluate the feasibility and preliminary effectiveness of *tailored care* interventions in subsets of patients with specific health needs.

Part 1. Assessing health needs of individuals with osteoarthritis

In this first part of the thesis the central theme concerns the identification of clinical relevant *health needs* of people with osteoarthritis in different stages of the disease. In current osteoarthritis-related research the most assessed and studied health needs are pain, stiffness/symptoms and limitations in activities, most often assessed by use of the Western Ontario and McMaster University Osteoarthritis index (WOMAC) scale. Measurement validity of assessment instruments is crucial to make clinically relevant decisions for our patients. So, to tailor an intervention to a patient's needs, valid outcome measures are necessary. The WOMAC scale has been criticized because of its inability to discriminate well enough between pain and physical functioning. Therefore, to set meaningful treatment goals a measurement instrument other than the WOMAC is needed. In **Chapter 2** we report on the validation of the Dutch language version of the Lower Extremity Functional Scale (LEFS) and discuss its ability to discriminate between pain and functioning.

A factor that might have an influence on the patients' health status is the presence or absence of joint-pain comorbidity (33;43). To date, the influence of joint pain elsewhere

has only been marginally studied in an osteoarthritis population (52). Individuals with pain elsewhere might benefit from more systemic therapy approaches than people with only knee osteoarthritis. In **Chapter 3** and **Chapter 4**, we investigate our hypothesis that patients with joint pain comorbidities might experience more deteriorated health status in people referred to secondary care for hip or knee osteoarthritis, in a cross-sectional and longitudinal study respectively.

Another group at risk for compromised health status might be patients awaiting elective total joint replacement surgery. Total joint replacement surgery is widely considered an effective and successful end-stage surgical procedure for relieving pain and improving functional status (53;54). However, a significant number of patients experience continuing pain and functional disability after joint replacement (40;41). Patients with osteoarthritis scheduled for an elective major total joint replacement already endured a lasting period of gradual functional decline (6;55). Since preoperative functioning is considered the strongest determinant of postoperative pain and functioning (56-61), it is important to understand whether prolonged wait times are associated with deterioration in the preoperative status. In a systematic review, **Chapter 5**, the impact of waiting on pain and functional status is studied and risk factors for further decline were identified.

Finally, as mentioned earlier, *health needs* comprises not only the health status of patients, but also the entire context of the health care system. As yet, little is known about the course of health care consumption in osteoarthritis and factors that influence the consumption of health care. Insight in factors predicting health care use over time, will allow health professionals to identify groups of patients who might be at risk for suboptimal use of the health care system. Therefore, in **Chapter 6** factors that could predict use or non-use of the health care system in people with early symptomatic hip or knee osteoarthritis were identified.

Part 2. Tailoring care to the needs of patients with advanced osteoarthritis

In the second part of this thesis the development, feasibility and preliminary effectiveness of non-pharmacological treatment for two subgroups of patients, namely patients with joint-pain comorbidities and frail elderly patients awaiting total joint replacement is described. Both patient groups comprise patients of higher age with multiple comorbidities warranting tailored care.

When developing an intervention for these patients, several authors recommend a systematic approach as this will most likely result in more effective interventions (62;63). Furthermore, the development of an evidence based treatment program should synthesize information from expert opinions, patient perspectives, and research evidence (64). In **Chapter 7** such a systematic and stepped method to develop a multi-disciplinary and multidimensional care intervention tailored to the needs of complex, individual patients is described; hand osteoarthritis was used as case example.

Self-management for people with generalized osteoarthritis

In **Chapter 8** we implement the recommendations for the development of a multi-disciplinary intervention for people suffering from GOA and we describe a randomized clinical trial protocol to assess its effectiveness. Considering the impact of having joint pain elsewhere on a patient's health status, this sub-group of osteoarthritis patients will particularly benefit from a tailored approach. However, before embarking on a costly and time consuming endeavor like a randomized clinical trial, the feasibility and preliminary effectiveness of a somewhat similar multi-disciplinary conservative intervention in this group of patients is discussed in **Chapter 9**. In this chapter the potential and feasibility of a multi-baseline design in the evaluation of osteoarthritis is discussed.

Preoperative exercise for frail elderly awaiting total joint replacement

Preoperative factors associated with delayed and disappointing recovery after total joint replacement are being of older age (57;59), having a greater number of co-morbidities and being more physically impaired before surgery (56-61). Therefore, it might be of value to target these individuals, aiming to enhance the postoperative functional recovery. In **Chapter 10**, we aim to investigate whether patients awaiting total joint replacement could benefit from a preoperative exercise program. In this review the therapeutic quality of interventions is taken into account and its influence on study outcomes is assessed. Finally, results of a randomized clinical pilot trial on the feasibility and preliminary effectiveness of a high-intensity exercise program specifically tailored to the patient group that is predisposed for delayed postoperative recovery is presented in **Chapter 11**.

Chapter 12 is a general discussion, in which the main results are presented and critically appraised. It also offers recommendations for clinical practice and future research.

RERERENCE LIST

- (1) Poos MJJC, Gijzen P, Gommer AM. Achtergronden en details bij cijfers uit huisartsenregistraties. 24 juni 2009 ed: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Bilthoven: RIVM 2009;24 juni 2009.
- (2) Pereira D, Peleteiro B, Araujo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. *Osteoarthritis Cartilage*. 2011;19(11):1270-85.
- (3) Gunther KP, Sturmer T, Sauerland S, Zeissig I, Sun Y, Kessler S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. *Ann Rheum Dis*. 1998;(12):717-23.
- (4) Forestier R, Francon A, Briole V, Genty C, Chevalier X, Richette P. Prevalence of generalized osteoarthritis in a population with knee osteoarthritis. *Joint Bone Spine*. 2011;(3):275-8.
- (5) Laberge MA, Baum T, Virayavanich W, Nardo L, Nevitt MC, Lynch J, et al. Obesity increases the prevalence and severity of focal knee abnormalities diagnosed using 3T MRI in middle-aged subjects-data from the Osteoarthritis Initiative. *Skeletal Radiol*. 2012;41(6):633-41.
- (6) van Dijk GM, Dekker J, Veenhof C, Van den Ende CH. Course of functional status and pain in osteoarthritis of the hip or knee: a systematic review of the literature. *Arthritis Rheum*. 2006;55(5):779-85.
- (7) Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum*. 1991;34(5):505-14.
- (8) Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. *Osteoarthritis Cartilage*. 2005;(9):769-81.
- (9) Kidd BL. Osteoarthritis and joint pain. *Pain*. 2006;(1-2):6-9.
- (10) Chapple CM, Nicholson H, Baxter GD, Abbott JH. Patient characteristics that predict progression of knee osteoarthritis: a systematic review of prognostic studies. *Arthritis Care Res (Hoboken)*. 2011;(8):1115-25.
- (11) Sinusas K. Osteoarthritis: diagnosis and treatment. *Am Fam Physician* 2012;(1):49-56.
- (12) Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011;(9783):2115-26.
- (13) Aurich M, Squires GR, Reiner A, Mollenhauer JA, Kuettner KE, Poole AR, et al. Differential matrix degradation and turnover in early cartilage lesions of human knee and ankle joints. *Arthritis Rheum*. 2005;(1):112-9.
- (14) Goldring MB. The role of the chondrocyte in osteoarthritis. *Arthritis Rheum*. 2000;(9):1916-26.
- (15) Burrage PS, Mix KS, Brinckerhoff CE. Matrix metalloproteinases: role in arthritis. *Front Biosci*. 2006;529-43.
- (16) Karsenty G. An aggrecanase and osteoarthritis. *N Engl J Med*. 2005;(5):522-3.
- (17) Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011;(9783):2115-26.

- (18) Conaghan PG. Structural correlates of osteoarthritis pain: lessons from magnetic resonance imaging. In: Felson DT, Schaible HG, Editors, editors. Pain in osteoarthritis. New Jersey: Wiley-Blackwell; 2009.
- (19) Sellam J, Berenbaum F. The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat Rev Rheumatol*. 2010;(11):625-35.
- (20) Sofat N, Ejindu V, Kiely P. What makes osteoarthritis painful? The evidence for local and central pain processing. *Rheumatology (Oxford)*. 2011;50(12):2157-65.
- (21) Frobell RB, Le Graverand MP, Buck R, Roos EM, Roos HP, Tamez-Pena J, et al. The acutely ACL injured knee assessed by MRI: changes in joint fluid, bone marrow lesions, and cartilage during the first year. *Osteoarthritis Cartilage*. 2009(2):161-7.
- (22) Felson DT, McLaughlin S, Goggins J, LaValley MP, Gale ME, Totterman S, et al. Bone marrow edema and its relation to progression of knee osteoarthritis. *Ann Intern Med*. 2003(5 Pt 1):330-6.
- (23) Felson DT, Chaisson CE, Hill CL, Totterman SM, Gale ME, Skinner KM, et al. The association of bone marrow lesions with pain in knee osteoarthritis. *Ann Intern Med*. 2001;(7):541-9.
- (24) Sowers MF, Hayes C, Jamadar D, Capul D, Lachance L, Jannausch M, et al. Magnetic resonance-detected subchondral bone marrow and cartilage defect characteristics associated with pain and X-ray-defined knee osteoarthritis. *Osteoarthritis Cartilage*. 2003;(6):387-93.
- (25) Baliki MN, Geha PY, Jabakhanji R, Harden N, Schnitzer TJ, Apkarian AV. A preliminary fMRI study of analgesic treatment in chronic back pain and knee osteoarthritis. *Mol Pain*. 2008;47.
- (26) Kulkarni B, Bentley DE, Elliott R, Julyan PJ, Boger E, Watson A, et al. Arthritic pain is processed in brain areas concerned with emotions and fear. *Arthritis Rheum*. 2007;(4):1345-54.
- (27) Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, et al. Understanding the pain experience in hip and knee osteoarthritis--an OARSI/OMERACT initiative. *Osteoarthritis Cartilage*. 2008;(4):415-22.
- (28) Dekker J, van Dijk GM, Veenhof C. Risk factors for functional decline in osteoarthritis of the hip or knee. *Curr Opin Rheumatol*. 2009;(5):520-4.
- (29) Pisters MF, Veenhof C, van Dijk GM, Heymans MW, Twisk JW, Dekker J. The course of limitations in activities over 5 years in patients with knee and hip osteoarthritis with moderate functional limitations: risk factors for future functional decline. *Osteoarthritis Cartilage*. 2012;20(6):503-10.
- (30) Felson DT, Gross KD, Nevitt MC, Yang M, Lane NE, Torner JC, et al. The effects of impaired joint position sense on the development and progression of pain and structural damage in knee osteoarthritis. *Arthritis Rheum*. 2009;(8):1070-6.
- (31) Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage*. 2008;16(2):137-62.
- (32) Association DO. Diagnostiek en behandeling van heup- en knieartrose. 2007 [cited 2011 05-10-2011]; Available from: http://www.cbo.nl/Downloads/363/r1_heup_knie_07.pdf [In Dutch]
- (33) Conaghan PG, Birrell F, Burke M, Cumming J, Dickson J, Dieppe P, et al. osteoarthritis: national clinical guideline for care and management in adults. 2008.

- (34) Smink AJ, Van den Ende CH, Vliet-Vlieland TP, Swierstra BA, Kortland JH, Bijlsma JW, et al. Beating Osteoarthritis: Development of a stepped care strategy to optimize utilization and timing of conservative treatment modalities for patients with hip or knee osteoarthritis. *Clin Rheumatol*. 2011;30(12):1623-9.
- (35) Dieppe P, Lim K, Lohmander S. Who should have knee joint replacement surgery for osteoarthritis? *Int J Rheum Dis*. 2011;(2):175-80.
- (36) Snijders GF, den Broeder AA, van Riel PL, Straten VH, de Man FH, van den Hoogen FH, et al. Evidence-based tailored conservative treatment of knee and hip osteoarthritis: between knowing and doing. *Scand J Rheumatol*. 2011;(3):225-31.
- (37) Zhang W, Doherty M, Arden N, Bannwarth B, Bijlsma J, Gunther KP, et al. EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis*. 2005;64(5):669-81.
- (38) Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis Part III: changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage*. 2010;18(4):476-99.
- (39) Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database Syst Rev*. 2009;(3):CD007912.
- (40) Hawker GA. Who, when, and why total joint replacement surgery? The patient's perspective. *Curr Opin Rheumatol*. 2006;(5):526-30.
- (41) Wylde V, Dieppe P, Hewlett S, Learmonth ID. Total knee replacement: is it really an effective procedure for all? *Knee*. 2007;(6):417-23.
- (42) Kovac SH, Saag KG, Curtis JR, Allison J. Association of health-related quality of life with dual use of prescription and over-the-counter nonsteroidal antiinflammatory drugs. *Arthritis Rheum*. 2008;(2):227-33.
- (43) Driban JB, Sitler MR, Barbe MF, Balasubramanian E. Is osteoarthritis a heterogeneous disease that can be stratified into subsets? *Clin Rheumatol*. 2010;(2):123-31.
- (44) Wright AA, Cook CE, Flynn TW, Baxter GD, Abbott JH. Predictors of response to physical therapy intervention in patients with primary hip osteoarthritis. *J Clin Epidemiol*. 2011;(4):510-24.
- (45) Veenhof C, Van den Ende CH, Dekker J, Kieke AJ, Oostendorp RA, Bijlsma JW. Which patients with osteoarthritis of hip and/or knee benefit most from behavioral graded activity? *Int J Behav Med*. 2007;(2):86-91.
- (46) Hoeksma HL, Dekker J, Runday HK, Breedveld FC, Van den Ende CH. Manual therapy in osteoarthritis of the hip: outcome in subgroups of patients. *Rheumatology (Oxford)*. 2005;(4):461-4.
- (47) Hunter DJ, Hellio Le Graverand-Gastineau MP. How close are we to having structure-modifying drugs available? *Rheum Dis Clin North Am*. 2008;(3):789-802.
- (48) Herrero-Beaumont G, Roman-Blas JA, Castaneda S, Jimenez SA. Primary osteoarthritis no longer primary: three subsets with distinct etiological, clinical, and therapeutic characteristics. *Semin Arthritis Rheum*. 2009;(2):71-80.

- (49) Knoop J, van der Leeden M, Thorstensson CA, Roorda LD, Lems WF, Knol DL, et al. Identification of phenotypes with different clinical outcomes in knee osteoarthritis: Data from the osteoarthritis initiative. *Arthritis Care Res (Hoboken)*. 2011.
- (50) Ruta DA, Duffy MC, Farquharson A, Young AM, Gilmour FB, McElduff SP. Determining priorities for change in primary care: the value of practice-based needs assessment. *Br J Gen Pract*. 1997;(419):353-7.
- (51) Stevens A, Gabbay J. Needs assessment needs assessment. *Health Trends*. 1991;(1):20-3.
- (52) Suri P, Morgenroth DC, Kwoh CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. *Arthritis Care Res(Hoboken)*. 2010;62(12):1715-23.
- (53) Ewald FC, Wright RJ, Poss R, Thomas WH, Mason MD, Sledge CB. Kinematic total knee arthroplasty: a 10- to 14-year prospective follow-up review. *J Arthroplasty*. 1999;(4):473-80.
- (54) Anderson JG, Wixson RL, Tsai D, Stulberg SD, Chang RW. Functional outcome and patient satisfaction in total knee patients over the age of 75. *J Arthroplasty*. 1996;(7):831-40.
- (55) Sharma L, Cahue S, Song J, Hayes K, Pai YC, Dunlop D. Physical functioning over three years in knee osteoarthritis: role of psychosocial, local mechanical, and neuromuscular factors. *Arthritis Rheum* 2003;(12):3359-70.
- (56) Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol*. 2002;(6):1273-9.
- (57) Kennedy LG, Newman JH, Ackroyd CE, Dieppe PA. When should we do knee replacements? *Knee*. 2003;(2):161-6.
- (58) Fitzgerald JD, Orav EJ, Lee TH, Marcantonio ER, Poss R, Goldman L, et al. Patient quality of life during the 12 months following joint replacement surgery. *Arthritis Rheum*. 2004;(1):100-9.
- (59) Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am*. 2004;(10):2179-86.
- (60) Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI, Arenaza JC, et al. Effect of patient characteristics on reported outcomes after total knee replacement. *Clin Rheumatol*. 2007;(1):112-9.
- (61) Zeni JA, Jr., Snyder-Mackler L. Preoperative predictors of persistent impairments during stair ascent and descent after total knee arthroplasty. *J Bone Joint Surg Am*. 2010;(5):1130-6.
- (62) Green LWK. *Health promotion planning: an Educational and Ecological Approach*. 3 ed. Mountain View, California: Mayfield Publishing Company; 1999.
- (63) Kok G, van den Borne B, Mullen PD. Effectiveness of health education and health promotion: meta-analyses of effect studies and determinants of effectiveness. *Patient Educ Couns*. 1997;(1):19-27.
- (64) Haynes RB, Sackett DL, Gray JM, Cook DJ, Guyatt GH. Transferring evidence from research into practice: 1. The role of clinical care research evidence in clinical decisions. *ACP J Club*. 1996;(3):A14-A16.
- (65) Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. *Health Promot Int*. 2006;21(3):245-55.

Part I

Assessing the needs of people with
osteoarthritis

2

The highly reliable, valid and responsive Dutch Lower Extremity Functional Scale is the preferred measure for functioning in hip or knee osteoarthritis: a validation study.

ABSTRACT

Background. The WOMAC is the most widely used self-report measure to evaluate physical functioning in hip or knee osteoarthritis, however its ability to discriminate pain and physical functioning (i.e. discriminate validity) has repeatedly been questioned. Little to no data is available on the discriminant validity of alternative questionnaires that measure the same construct, for instance the Hip and Knee Osteoarthritis Outcome Score (HOOS and KOOS, respectively) and the Lower Extremity Function Scale (LEFS). Therefore, we translated the LEFS to Dutch and studied its psychometric properties (i.e. validity, reliability and responsiveness). In addition, we assessed the discriminant validity of the LEFS, HOOS and KOOS.

Methods. After translation with a forward/backward protocol, 401 individuals with hip or knee osteoarthritis completed the LEFS, HOOS/KOOS, SF-36, Hospital Anxiety and Depression Scale and Checklist Individual Strength questionnaires. To assess reliability and responsiveness, a sample of 106 and 108 patients completed a comparable set of questionnaires within 3 weeks and 3 months, respectively. Feasibility, validity, reliability and responsiveness were evaluated. Discriminant validity of the LEFS, HOOS and KOOS was examined by contrasting the scales' correlations with the physical functioning subscale of the SF-36 with the scales' correlations with the bodily pain subscale of the SF-36.

Results. The Dutch version of the LEFS was feasible, had good internal consistency (0.96), good reliability (ICC=0.86), good construct and discriminant validity, and showed no floor or ceiling effects. The minimal detectable change (MDC90) was ten points. Area under the receiver operating characteristic curve (AUC) analyses revealed good (AUC=0.76) and fair (AUC=0.63) responsiveness for the LEFS in improved and worsened patients, respectively. Discriminant validity for pain was apparent for the LEFS ($p<0.01$), but not for the HOOS and KOOS ($p=0.21$ and $p=0.20$, respectively).

Conclusions. Considering the LEFS' good psychometric qualities and ability to discriminate between pain and functioning, we recommend the LEFS as the outcome measure of choice to assess self-reported physical functioning in individuals with hip or knee osteoarthritis.

Hoozeboom TJ, de Bie RA, den Broeder AA, van den Ende CHM.

The highly reliable, valid and responsive Dutch Lower Extremity Functional Scale is the preferred measure for functioning in hip or knee osteoarthritis: a validation study.

BMC Musculoskeletal Disorders 2012;13(1),117.

BACKGROUND

Numerous self-report measures on physical function are available for the evaluation of patients with hip or knee osteoarthritis [1]. Among those, the licensed for use Western Ontario and McMaster University Osteoarthritis Index (WOMAC) [2] is the most widely used [3]. It is recommended by the Osteoarthritis Research Society for use in clinical trials in patients with hip or knee osteoarthritis to measure pain and disability [4]. However, consensus statements consistently advocate that pain and physical function must be measured independently [3, 5]. A solid body of evidence demonstrates that the WOMAC-PF (Physical Function subscale) is unable to discriminate between pain and function [6-9]. Recently, three new license free self-report measures to determine functioning in patients with osteoarthritis have become available; the Hip Osteoarthritis Outcome Score (HOOS) [10], the Knee Osteoarthritis Outcome Score (KOOS) [11] and the Lower Extremity Function Scale (LEFS) [12]. One of those new measures, the LEFS, showed promise as a competitive alternative to the WOMAC-PF, as the LEFS can differentiate pain and functioning [13] and detect changes in functional status in the period immediately following surgery [14]. Moreover, the LEFS has excellent test-retest reliability, internal consistency and construct validity [12, 13, 15]. To date, it remains to be seen, whether the physical function scales of the HOOS and KOOS can discriminate between pain and physical function [10, 11, 16].

Since the LEFS is currently not available in Dutch, the primary purpose of this study was to evaluate the psychometric qualities of the Dutch LEFS in people with hip or knee osteoarthritis. Our secondary objective was to assess the discriminant validity for pain of the physical function subscale of the HOOS and KOOS and the LEFS.

METHODS

First the English version of the LEFS was translated into Dutch according to a standardized procedure described by Beaton *et al* [17], and secondly it was tested for psychometric quality by use of prospective data.

Procedure of translation

The translation procedure consisted of four steps. First, two persons translated independently of each other the English version of the LEFS into Dutch (forward translation) (T1 & T2); one translator (TJH) had a medical background and was familiar with the concepts of the questionnaire and the other (VvS) was a certified translator without a medical background. Both were native speakers. Based on a consensus meeting one final

version (T-12) was formed. Second, two bilingual persons (T3 & T4) translated the T-12 questionnaire back into English (BT1 & BT2), to guarantee a consistent translation of the questionnaire. Both translators (PA & DKJ) were unfamiliar with the original questionnaire, the concepts of the questionnaire, and had no medical background. DKJ is also a certified translator. Third, an expert meeting was organised in which all translators, two health professionals (CKS, ML), a methodologist (CHMvE) and two language experts participated. During this meeting all versions of questionnaires (T1, T2, T-12, BT1, BT2) were combined and consensus on semantic, idiomatic, experiential and conceptual equivalence was reached resulting in a pre-final version of the questionnaire. The developers of the original questionnaire approved all previous steps and the final version. Finally, the pre-final version was presented in a group of 33 patients (20 women and 13 men; age (SD): 63 (13) years) to explore the clarity of the questionnaire. All patients were asked whether they understood the items and whether they could interpret the questionnaires correctly. Also, the time needed to complete the questionnaire was timed. The findings were discussed among the translators, resulting in only minor changes to the final Dutch version of the LEFS. Mean completion time was 3.5 (SD = 1.5) minutes. For the final version of the Dutch LEFS see Appendix 1.

Patients and procedure

Individuals (≥ 18 years) diagnosed with hip or knee osteoarthritis (inclusion period June till October 2009) by an orthopaedic surgeon in the Sint Maartenskliniek hospital Nijmegen were eligible. People reporting concurrent rheumatoid arthritis, fibromyalgia or psoriatic arthritis, were excluded. Written materials were sent by mail: this included an information letter, an informed consent form, the questionnaires and a return envelope. At baseline, all patients completed four questionnaires, the LEFS, the HOOS or KOOS (depending on index joint), the SF-36 and the Hospital Anxiety and Depression Scale (HADS). A reminder was sent to those patients who did not respond within three weeks, to ensure a high response rate. One-hundred and twenty participants were sent a follow-up questionnaire to evaluate test-retest reliability (within 3 weeks) and another 120 participants were sent a follow-up questionnaire to evaluate responsiveness (after 3 months); as 100 participants were deemed sufficient [13]. By use of random numbers the 240 patients were selected to either the reliability or responsiveness study. Both follow-up mailings consisted of three questionnaires (LEFS, HOOS or KOOS, and the SF-36) and a global perceived effect question. For test-retest reliability, we considered a time interval of 3 weeks to be appropriate for the current population. For responsiveness, we deemed a period up to 3 months long enough to allow for improvement and brief enough to minimize the risk of a response shift [18, 19]. The study was approved by the Institutional Review Board of the University Medical Centre Nijmegen (ID: 2009/20).

Measures

The LEFS is a 20-item condition-specific questionnaire designed to be applicable to individuals with musculoskeletal conditions of the lower extremity [12]. Each item of the LEFS scores on a 5-point scale ranging from 0 to 4 points. When scoring the LEFS, up to 4 missing item responses are permitted, for more detailed information see Stratford et al (2005) [20]. Accordingly, LEFS scores range from 0 to 80 points, with higher scores representing higher levels of functioning.

The HOOS and the KOOS include five subscales: Pain, other Symptoms, Function in Daily living (ADL), Function in Sport and Recreation (Sport/Rec), and hip/knee-related quality of life (QoL). Standardized response options are given (5-point Likert scale) and each question is scored from 0 to 4 points. Subsequently, a normalized score (100 indicating no symptoms and 0 indicating extreme symptoms) is calculated for each subscale. The Dutch HOOS and KOOS have good internal consistency, construct validity, no floor and ceiling effects and have been found to be reliable [10, 11]. Both the HOOS and KOOS questionnaires include the WOMAC osteoarthritis-index in its complete and original format (with permission, www.koos.nu).

The SF-36 is a generic health status questionnaire which contains 36 items [21]. It measures eight major attributes (bodily pain; physical function; social function; role limitations because of physical problems; role limitations because of emotional problems; mental health; vitality; general health perceptions). It is widely used, reliable, validated into Dutch and is easy to complete. Higher scores indicate better health [22].

The Hospital Anxiety and Depression Scale (HADS) is a 14-item scale designed to detect anxiety and depression, independent of somatic symptoms [23]. It consists of two 7-item subscales measuring depression and anxiety on a 4-point response scale (from 0, no symptoms, to 3, maximum symptoms), with possible scores for each subscale ranging from 0 to 21. HADS is a valid and reliable screening instrument for detecting mood disorder in people with osteoarthritis [24, 25]. Higher scores indicate higher levels of disorder.

Fatigue is measured with the 8-itemed “Subjective Fatigue” subscale of the Checklist Individual Strength (CIS) [26]. The outcomes per question are given in a 7-point scale, ranging from the statement ‘totally right’ to the statement ‘totally wrong’. The total score is counted in points with a range of 1-7 per question and a total score range of 8-56 points. The CIS is a sensitive instrument with good discriminating power and reliability [26].

The external criterion for distinguishing between improved and unimproved subjects was a 7-point global perceived effect (GPE) scale. The categories of improvement included the following: completely recovered, much improved, slightly improved, not changed, slightly worse, much worse, and vastly worsened.

Statistical analyses

Descriptive statistics were used to describe the study population and the number of missing

values. Data symmetry was tested by use of visual inspection of the data distribution plotted by histograms. Psychometric qualities of the LEFS were expressed by floor- and ceiling effects, internal consistency, test-retest reliability, minimally detectable change, construct validity, discriminant validity and responsiveness.

Floor and ceiling effects

Floor and ceiling effects were determined by calculating the number of individuals that obtained the lowest (0) or highest (80) scores possible and were considered present if more than 15% of the participants achieved the highest or lowest score [27].

Internal consistency and dimensionality

Internal consistency – an indicator for the homogeneity of a questionnaire - was assessed with Cronbach's alpha and 95% confidence intervals (95% CI's). Internal consistency is considered good when Cronbach's alpha lies between 0.7 and 0.9 [28]. Dimensionality was assessed by performing principal component factor analysis with loading coefficient absolute value suppression at 0.40 on the LEFS, KOOS-PF and HOOS-PF to determine if the individual items loaded on a single factor. Factor extraction had three requirements: scree plot point of inflection at the second Eigenvalue, Eigenvalue cut-off >1.0, and $\geq 10\%$ variance [29].

Reliability and minimal detectable change

Reliability concerns the degree to which the results of measurement are consistent across repeated measurements [28]. Test-retest reliability of the Dutch LEFS was determined by means of Intraclass Correlation Coefficients (ICCs) (two-way random effects model absolute agreement) and Bland and Altman plots [30]. The ICC(2,1) equals variance between patients divided by variance between patients plus variance between measurements plus error variance. The value of the ICC ranges from 0 to 1, where one represents perfect reliability of the measurement. Consequently, to quantify the reliability of the LEFS scores we determined the standard error of measurement ($SEM = SD[\sqrt{1 - ICC}]$). The SEM is a representation of measurement error expressed in the same units as the original measurement. We quantified the minimal detectable change at the 90% and 95% confidence level (MDC_{90} and MDC_{95}) by multiplying the point estimate of the SEM, the square root of 2 (to account for the error associated with repeated measurements), and the z score of 1.65 or 1.96 (resp. 90% or 95% confidence level); formula $MDC_{90} = SEM * 1.65 * \sqrt{2}$ and $MDC_{95} = SEM * 1.96 * \sqrt{2}$ [31].

Validity

Construct validity reflects the extent to which a particular measure consistently relates to other measures with theoretically derived hypotheses for the constructs that are being measured [28]. To evaluate the construct validity of the LEFS, we formulated a set of 16

hypotheses (eight for knee osteoarthritis and eight for hip osteoarthritis) about the expected magnitude and direction of relationships between the LEFS and other instruments. If 75% or more of the arbitrarily set number of 16 hypotheses were confirmed we defined the construct validity of the LEFS as good [32, 33].

Discriminant validity was examined for the LEFS and the physical function subscale of the HOOS and KOOS, by contrasting its correlation with the PF subscale of the SF-36 with its correlation with the bodily pain subscale of the SF-36. Meng *et al*'s test for dependent data was used to evaluate the differences between those correlations [34].

Responsiveness

We studied the responsiveness of the LEFS and the WOMAC-PF extracted from the HOOS-PF and KOOS-PF) in a combined hip and knee group, as only a very small number of patients reported clinically important change, thus not allowing to study the responsiveness of the HOOS and KOOS separately. As yet, a variety of responsiveness statistics is available. However, it is not yet known which of these statistics is better for assessing responsiveness [35] we utilized three different analyses. First we determined the Responsiveness Ratio of Guyatt (GRI: average change of recovered patients (GPE = 1-2) / SD of average change of stable patients (GPE = 3-5)). If the responsiveness ratio is larger than 1, the mean change score in clinically improved patients exceeds the measurement error and the instrument may be considered to be responsive, to an extent that is proportional to the magnitude of the responsiveness ratio [36, 37]. Second, we determined the Standardized Response Mean (SRM: average score change / SD of score change). By use of the modified Jackknife testing, we assessed differences in SRM statistically [38]. Third, we calculated Receiver operating characteristic curves (ROC) for the improved subjects and for the worsened subjects using the change scores of the questionnaires and the patients' ratings of change [39]. The patients' rating of change was dichotomized to identify those subjects who experienced a clinically meaningful reduction of symptoms. Important change was defined as 'Much Improvement (GPE = 1-2)' or 'Much Decline (GPE = 6-7)'. Consequently, we computed the area under the curve (AUC). An AUC of 1.0 indicates perfect discrimination, whereas an AUC of 0.50 indicates no performance better than chance.

RESULTS

Four-hundred and one individuals returned the baseline questionnaire in the study (response rate 82%). After the baseline questionnaire, 121 participants received a follow-up mailing to evaluate test-retest reliability (106 responded (88%)) and 125 participants received a

Table 1. Patient and disease characteristics at baseline from the total group and the

Baseline sample		
<i>N</i>	401	
Age, mean (SD)	58 (13)	
Female, ♀%	231 (58%)	
Index joint, knee	284 (71%)	
BMI, median (IQR)	26 (24-29)	
Education, higher [†]	77 (20%)	
Duration complaints, >5 years [†]	144 (61%)	
Co-morbidities, yes	192 (52%)	
Baseline data		
	Hip group	Knee Group
<i>N</i>	117	284
LEFS	36.0 (16.6)	39.6 (14.1)
HOOS/KOOS		
Pain, mean (SD)	44.6 (21.3)	49.9 (20.0)
Symptoms, mean (SD)	41.6 (21.4)	54.6 (19.3)
Daily living, mean (SD)	47.5 (21.9)	56.5 (21.7)
Sports & recreation, median (IQR)	25 (6-38)	10 (0-30)
Quality of life, median (IQR)	25 (13-38)	25 (13-44)
SF36		
Pain, mean (SD)	35.6 (8.9)	37.3 (8.2)
Functioning, mean (SD)	31.8 (9.8)	32.8 (9.2)
HADS (total), median (IQR)	9 (5-16)	9 (4-15)
CIS, mean (SD)	34.8 (12.6)	31.0 (13.6)

* Figures are numbers and percentages in the brackets unless indicated differently. [†]Cut-off arbitrarily chosen.

follow-up mailing to evaluate the responsiveness (112 responded (90%)). Patient characteristics at baseline and follow-up are presented in Table 1.

The majority of patients (86%) had less than three missing values. The proportion of missing values in the LEFS questionnaire (4%) was slightly less than the proportion of missing values in the KOOS (5%) and the HOOS (8%) questionnaires. The item 'getting in or out of bath' had the highest number of missing values in each of the questionnaires; 5% in the HOOS, 7% in the LEFS and 10% in the KOOS.

Floor and ceiling effects

None of the 401 participants reported the lowest possible score whereas one patient (0.26%) reported the highest functional level implying that the Dutch LEFS has no floor or ceiling effects. In addition, the distribution of the LEFS was symmetrical.

follow-up data from the reliability and the responsiveness sample (continuation Table 1).*

Reliability sample (≤3 weeks)		Responsiveness sample (≥3 months)	
106		112	
61 (11)		58 (11)	
65 (63%)		59 (54%)	
81 (76%)		81 (72%)	
26 (24-30)		27 (24-30)	
25 (25%)		27 (25%)	
72 (63%)		59 (58%)	
52 (50%)		63 (72%)	
Follow-up data			
Hip group 25	Knee Group 81	Hip group 31	Knee Group 81
29.8 (13.5)	36.4 (13.7)	34.3 (17.0)	37.5 (15.0)
41.7 (21.1)	46.5 (19.7)	43.6 (23.7)	53.8 (19.3)
40.2 (20.3)	52.5 (20.5)	39.7 (23.5)	55.8 (19.6)
43.8 (21.0)	54.2 (20.3)	46.5 (24.0)	57.8 (21.4)
13 (6-25)	5 (0-25)	19 (6-31)	10 (0-25)
31 (19-38)	25 (13-38)	25 (13-38)	31 (19-44)
35.6 (6.5)	38.6 (7.3)	37.0 (8.0)	39.9 (8.3)
29.8 (8.5)	32.5 (8.2)	31.0 (10.0)	33.4 (9.5)
N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A

Abbreviations: BMI = Body Mass Index, IQR = Inter Quartile Range, SD = Standard Deviation.

Internal validity and factorial structure

The internal consistency for the total group of patients (n = 401) reached a Cronbach's alpha of 0.96 (lower limit (LL) 95%-CI: 0.95) for the 20 items. For the hip and knee osteoarthritis group Cronbach's alpha reached 0.97 (LL 95%-CI: 0.96) and 0.95 (LL 95%-CI: 0.94), respectively. Within-scale principal component factor analysis revealed that all items included in the LEFS, KOOS-PF and HOOS-PF loaded on a single major factor (Table.2).

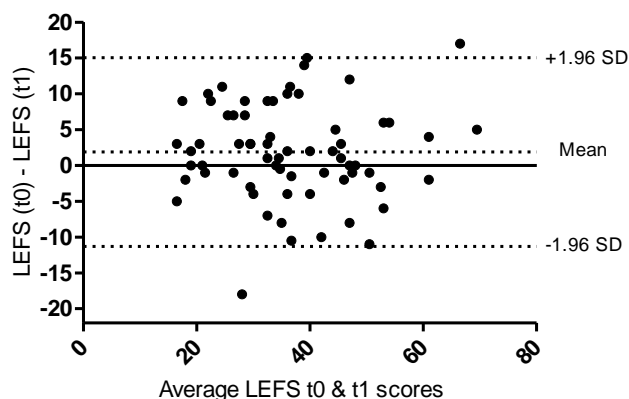
Table 2. Factor Analysis: Variance Explained for the LEFS, KOOS and HOOS.*

Factor	Initial Eigenvalues		
	Total	% of variance	Cumulative %
LEFS			
Factor 1	10.84	54.18	54.18
Factor 2	1.87	9.33	63.51
Factor 3-20	0.96 - 0.06	36.49	100.00
KOOS			
Factor 1	9.76	57.42	57.42
Factor 2	1.46	8.61	66.03
Factor 3-17	0.86 - 0.05	33.97	100.00
HOOS			
Factor 1	9.69	57.01	57.01
Factor 2	1.50	8.83	65.84
Factor 3	1.02	5.98	71.82
Factor 4	0.94 - 0.07	28.18	100.00

* A principal component factor analysis with loading coefficient absolute value suppression at 0.40

Reliability and minimal detectable change

Within three weeks after the baseline questionnaire, five individuals improved (5%) (GPE = 1-2), three worsened (3%) (GPE = 6-7) and the majority (92%) remained stable (GPE = 3-5). Two-way random effects ANOVA demonstrated that the ICC of the Dutch LEFS questionnaire for the total group (n = 106) was 0.86. For the knee group (n=81) and the hip group (n=25) the ICC was 0.87 and 0.78, respectively. The standard error of measurement was 4.4 points. The MDC₉₀ and MDC₉₅ of the LEFS questionnaire was 10 points and 12 points, respectively.

Figure 1. Bland & Altman plot.

The Bland-Altman plot (Figure 1) shows that the mean difference between the two applications of the LEFS was 1.87 points (95%-CI 0.22 to 3.52). The limits of agreement (mean \pm 1.96 SD) ranged from -11.56 to 15.30 points.

Validity

Thirteen of the 16 predefined hypotheses to determine the construct validity were confirmed (81%) (Table 3 and 4). The following three hypotheses could not be confirmed. In the hip group we found a correlation of 0.55 between LEFS and CIS scores, which was higher than the predefined cut-off of 0.5. In the knee group we found that the duration of complaints did not influence the LEFS scores and that education level (primary, secondary or higher education) did influence the LEFS scores.

Table 3. Predefined hypotheses and the confirmation or rejection of the hypotheses for hip OA (n = 117) and knee OA (n = 284).

	Hip OA	Knee OA
There is a moderate correlation ($r > 0.6$) between LEFS scores and HOOS/KOOS-PF subscale.	Yes	Yes
The correlation between the LEFS and the HOOS/KOOS-PF subscale is higher than the correlation between the LEFS and the other subscales of the HOOS/KOOS.	Yes	Yes
The correlation between LEFS and HADS scores is low ($r < 0.5$).	Yes	Yes
The correlation between LEFS and CIS scores is low ($r < 0.5$).	No	Yes
Patients with multiple painful lower limb joints demonstrate lower LEFS scores than patients with pain in a single joint.	Yes	Yes
Patients with complaints less than 5 years demonstrate higher LEFS scores than patients with complaints over 5 years.	Yes	No
The participants' education level (primary, secondary or higher education) does not influence the LEFS scores.	Yes	No
Sociodemographic characteristics, such as sex, doing volunteer work, income, type of health insurance, and family status, are unrelated ($r < 0.2$) to the LEFS score:	Yes	Yes
	7/8 (88%)	6/8 (75%)

Abbreviations: CIS = Checklist Individuals Strength, HADS = Hospital Anxiety and Depression Questionnaire, HOOS = Hip Osteoarthritis and Outcome Score, KOOS = Knee Osteoarthritis and Outcome Score, LEFS = Lower Extremity Functional Scale, OA = osteoarthritis, PF = physical function.

Meng *et al*'s test demonstrated that the association of the LEFS with the SF-36 subscale pain differed significantly with the SF-36 subscale physical functioning (Table 4), indicating that the LEFS has discriminant validity for pain ($p < 0.01$). We found no significant differences between the association with SF-36 subscale's pain and physical functioning and the HOOS-PF (r (95%-CI) = 0.64 (0.51 - 0.74) and 0.71 (0.60 - 0.79), $p =$

0.21) and the KOOS-PF (0.69 (0.62 - 0.75) and 0.73 (0.67 - 0.79), $p = 0.20$, respectively), indicating that both questionnaires do not discriminate between pain and physical functioning.

Table 4. Correlation data for the hip OA (n = 117) and knee OA (n = 284) groups at baseline.

	Correlation with LEFS	
	Hip group	Knee group
	r (95%-CI)	r (95%-CI)
HOOS/KOOS		
Pain	0.71 (0.60 – 0.79)	0.65 (0.58 – 0.72)
Symptoms	0.56 (0.41 – 0.67)	0.42 (0.32 – 0.52)
Daily living	0.78 (0.69 – 0.84)	0.78 (0.73 – 0.83)
Sports and recreation	0.74 (0.64 – 0.82)	0.63 (0.55 – 0.70)
Quality of life	0.72 (0.62 – 0.80)	0.60 (0.52 – 0.67)
SF36		
Pain	0.51 (0.35 – 0.64)	0.66 (0.59 – 0.72)
Functioning	0.82 (0.75 – 0.88)	0.81 (0.77 – 0.85)
HADS		
Anxiety	-0.25 [†] (-0.42 – -0.07)	-0.10 (-0.22 – 0.02)
Depression	-0.30 [†] (-0.46 – -0.12)	-0.28 (-0.38 – -0.16)
CIS	-0.55 [†] (-0.67 – -0.41)	-0.47 (-0.56 – -0.37)

[†]Due to the inverted scales, the correct interpretation of negative associations is that more anxiety/depression/fatigue is associated with worse function. **Abbreviations:** CIS = Checklist Individuals Strength, HADS = Hospital Anxiety and Depression Questionnaire, HOOS = Hip Osteoarthritis and Outcome Score, IQR = Inter-quartile range, KOOS = Knee Osteoarthritis and Outcome Score, LEFS = Lower Extremity Functional Scale, OA = osteoarthritis.

Responsiveness

Seven people (7%) reported relevant improvements in function (GPE = 1-2), nine people reported relevant worsening (8%) (GPE = 6-7) and the majority remained stable (85%) (GPE = 3-5). Responsiveness Ratio of the LEFS was 1.49, close to the outcomes of WOMAC-PF (1.20) and SF36-PF (1.22) (Table 5).

Modified Jackknife testing demonstrated no statistical differences between the SRM for the LEFS (0.13) compared with the SRM of the WOMAC (SRM = 0.02, $p = 0.45$) and SF-36 (SRM = 0.00, $p = 0.36$). ROC curve analysis revealed that for improved patients the AUC was 0.76 (95% CI: 0.49 - 1.00) for the LEFS, 0.71 (95% CI: 0.45 - 0.98) for the WOMAC-PF (extracted from the HOOS-PF and KOOS-PF) and 0.68 (0.44 - 0.93) for the SF36-PF. For worsened patients the AUC was 0.63 (95% CI: 0.42 - 0.83) for the LEFS, 0.56 (0.34 - 0.78) for the WOMAC-PF and 0.56 (0.35 - 0.78) for the SF36-PF.

Table 5. Responsiveness ratio (GRI).

	Mean change	SD (baseline)	SD (change)	Mean change (GPE = 1-2)	SD mean change (GPE = 3-5)	SRM	GRI
LEFS	1.42	14.41	11.25	12.50	8.38	0.13	1.49
WOMAC- PF	0.19	14.84	8.86	9.66	8.06	0.02	1.20
SF36-PF	0.00	22.73	13.68	14.44	11.86	0.00	1.22

Abbreviations: GPE = Global perceived effect, GRI = Guyatt Responder Ratio, LEFS = Lower Extremity Functional Scale, SD = Standard Deviation, SF-36-PF = Short Form 36 - Physical Function scale, SRM = Standardized Response Mean, WOMAC = Western Ontario and McMasters University Osteoarthritis Index - Physical Function scale.

DISCUSSION

The primary objective of this study was to create a reliable and valid Dutch version of LEFS by translation and adaptation. No difficulties were encountered in the translation phase of the study; the structure of the original LEFS was not altered and all items were maintained. Moreover, participants reported no problems in the administration of the questionnaire. Considering the results of this validation study, we deemed the Dutch version of the LEFS to be an internally consistent, uni-dimensional, highly reliable and valid questionnaire to determine lower extremity functioning in patients with hip or knee osteoarthritis. Finally, the LEFS revealed good responsiveness by detecting improvement in patient GPE; however this finding should be interpreted with caution, given the small proportion of patient to actually report clinically relevant functional improvement. For our secondary objective, we were unable to demonstrate that the HOOS-PF and KOOS-PF subscales are able to discriminate between pain and physical function.

Construct validity of the Dutch version of the LEFS was good as most of the pre-formulated hypotheses were met. Three of the 16 hypotheses could however not be confirmed. First, in the hip group, the correlation between the lower extremity functioning (LEFS) and fatigue (CIS) was over 0.5 in the hip group, however similar correlations were found for HOOS-PF ($r = 0.55$) and SF-36 PF ($r = 0.50$). As comparative measures also demonstrate such a relation, fatigue might have a stronger relation with functioning than previously thought [10, 15]. An important difference with previous studies is that we investigated fatigue with a fatigue-specific questionnaire in contrast to others that used the vitality scale of the SF-36 [10, 15]. Second, participants with knee symptoms for less than five years did not report significantly less symptoms than patients with symptoms for over 5 years. Again this finding was also found for the KOOS-PF ($p = 0.90$) and the SF-36 PF ($p = 0.75$). These findings, could however, be biased by a phenomenon called response shift,

which could have resulted in an underreporting of functional disabilities in the group with the longest duration of complaints [40]. Third and final, in the knee group we found that participants' education level (primary, secondary or higher education) did influence the LEFS scores. It would be undesirable if LEFS scores were influenced by education level, as this would indicate that the LEFS is difficult to interpret. Further scrutiny of this finding indicates that patients with knee symptoms who enjoyed a higher education reported less symptoms than patients without or only primary education ($p = 0.02$); also when adjusted for age, sex, BMI, co-morbidities, duration of complaints and being employed. Yet again, this finding was also found for the KOOS-PF ($p = 0.04$), but not for the SF-36 PF subscale ($p = 0.08$). Our findings are in contrast to a previous study that addressed the relation between the LEFS scores (Italian version) and education levels. This discrepancy can possibly be explained by the different format of the Italian version; an interview-format instead of a self-reported questionnaire [15]. It would be of interest to further elucidate this relation in other studies.

Although the responsiveness of the Dutch LEFS was good and superior to the WOMAC-PF and SF36-PF, compared to Italian validation study by Cacchio *et al* (2010) (AUC=0.86) it was somewhat low [15]. On the other hand, the psychometric properties of the Dutch LEFS (i.e. Cronbach's alpha [12, 15], reproducibility [12, 13, 15] and validity [12, 13]) were comparable to the findings of previous validation studies. Our results regarding the responsiveness of the LEFS, WOMAC-PF and SF36-PF, should be interpreted with caution. Given the small number of patients reporting clinically relevant change which may have impacted for example the magnitude of the SRM, the point estimates might be spurious. Future (intervention) studies should further investigate the responsiveness of the Dutch LEFS.

The lack of discriminant validity for the WOMAC-PF has been demonstrated in numerous occasions [6-9, 41, 42]. Therefore, the greater discriminant validity of the LEFS compared to the WOMAC-PF [13, 14] was one of the foremost reasons to translate and adapt the LEFS to the Dutch language. In our study we compared the LEFS questionnaire to the HOOS-PF and KOOS-PF subscales. As the physical function subscale of the HOOS and KOOS are very similar to the WOMAC-PF, these subscales are also at great risk for lacking discriminant validity. Our results indicate that the LEFS, but not the KOOS-PF and the HOOS-PF, could discriminate from pain measures, that is, KOOS-PF and HOOS-PF did not show a statistically higher correlation with the PF subscale than with the bodily pain subscale of the SF-36, whereas the LEFS did. As far as we know, we are the first to also demonstrate the lack of discriminant validity in the (Dutch version of the) HOOS and KOOS subscales, as in those particular validation studies only SF-36 subscales other than the bodily pain subscale were examined [10, 11, 16].

A limitation of our study is that we recruited only individuals with hip and knee osteoarthritis. Originally the LEFS has been developed as a measure that could be used for all kinds of conditions of the lower extremity [12]. The exclusion of other condition

hampers the generalizability of our findings to other complaints of the lower extremity. We did however evaluate the LEFS ability to differentiate between patients with and without additional lower extremity pain co-morbidities, which demonstrated a linear association between the number of lower extremity joint pain co-morbidities and LEFS scores. The latter analysis showed promise that the Dutch version of the LEFS is also able to detect functional disabilities in patients with other symptoms than just hip and knee osteoarthritis. Another limitation of this study is that we did not assess the association between the Dutch version of the LEFS and a set of performance measures to determine the convergent validity. Future studies should investigate this association. A third limitation, the Cronbach's Alpha value surpassed the cut-off value of 0.90 indicating item redundancy. However, due to the magnitude of our study sample and relatively high number of items this figure might have been inflated [43]. Finally, we studied the construct validity of the LEFS by testing hypotheses according to prespecified cut-off values; however cut-off values are often too rigid by their dichotomous (true/false) nature. Future studies should consider using the lower or upper bound of the 95% confidence interval of an association.

CONCLUSIONS

We found that the Dutch version of the LEFS has no floor and ceiling effects, good internal consistency, reliability, construct validity and responsiveness. Moreover, the Dutch LEFS demonstrated discriminant validity for pain, as it was able to discriminate between pain and physical functioning, whereas both the HOOS-PF and KOOS-PF did not. Therefore, we recommend the use of the Dutch LEFS as an outcome measure for physical functioning in patients with hip and/or knee osteoarthritis.

REFERENCE LIST

- (1) Veenhof C, Bijlsma JW, van den Ende CH, van Dijk GM, Pisters MF, Dekker J: Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature. *Arthritis Rheum* 2006, 55(3):480-492.
- (2) Bellamy N: WOMAC osteoarthritis index. A user's guide. In. London, Ontario, Canada: University of Western Ontario; 1995.
- (3) Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, Altman R, Brandt K, Dougados M, Lequesne M: Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997, 24(4):799-802.
- (4) Hochberg MC, Altman RD, Brandt KD, Moskowitz RW: Design and conduct of clinical trials in osteoarthritis: preliminary recommendations from a task force of the Osteoarthritis Research Society. *J Rheumatol* 1997, 24(4):792-794.
- (5) Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N *et al*: Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005, 113(1-2):9-19.
- (6) Soderman P, Malchau H: Validity and reliability of Swedish WOMAC osteoarthritis index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). *Acta Orthop Scand* 2000, 71(1):39-46.
- (7) Soderman P, Malchau H: Is the Harris hip score system useful to study the outcome of total hip replacement? *Clin Orthop Relat Res* 2001(384):189-197.
- (8) Terwee CB, van der Slikke RM, van Lummel RC, Benink RJ, Meijers WG, de Vet HC: Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. *J Clin Epidemiol* 2006, 59(7):724-731.
- (9) Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JJ: Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index. *Clin Rheumatol* 2002, 21(6):466-471.
- (10) de Groot IB, Reijman M, Terwee CB, Bierma-Zeinstra SM, Favejee M, Roos EM, Verhaar JA: Validation of the Dutch version of the Hip disability and Osteoarthritis Outcome Score. *Osteoarthritis Cartilage* 2007, 15(1):104-109.
- (11) de Groot IB, Favejee MM, Reijman M, Verhaar JA, Terwee CB: The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes* 2008, 6:16.
- (12) Binkley JM, Stratford PW, Lott SA, Riddle DL: The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. *Phys Ther* 1999, 79(4):371-383.
- (13) Pua YH, Cowan SM, Wrigley TV, Bennell KL: The Lower Extremity Functional Scale could be an alternative to the Western Ontario and McMaster Universities Osteoarthritis Index physical function scale. *J Clin Epidemiol* 2009, 62(10):1103-1111.

- (14) Stratford PW, Kennedy DM, Hanna SE: Condition-specific Western Ontario McMaster Osteoarthritis Index was not superior to region-specific Lower Extremity Functional Scale at detecting change. *J Clin Epidemiol* 2004, 57(10):1025-1032.
- (15) Cacchio A, De Blasis E, Necozone S, Rosa F, Riddle DL, di Orio F, De Blasis D, Santilli V: The Italian version of the lower extremity functional scale was reliable, valid, and responsive. *J Clin Epidemiol* 2010, 63(5):550-557.
- (16) Roos EM, Toksvig-Larsen S: Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes* 2003, 1:17.
- (17) Beaton DE, Bombardier C, Guillemin F, Ferraz MB: Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000, 25(24):3186-3191.
- (18) Sprangers MA, Schwartz CE: Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med* 1999, 48(11):1507-1515.
- (19) Robling M, Hood K: Response shift, responsiveness or recall bias? *Br J Gen Pract* 2002, 52(480):585.
- (20) Stratford PW, Hart DL, Binkley JM, Kennedy DM, Alcock GK, Hanna SE: Interpreting Lower Extremity Functional Status Scores. *Physiotherapy Canada* 2005, 57(2):9.
- (21) Ware JE, Jr., Sherbourne CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *MedCare* 1992, 30(6):473-483.
- (22) Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A, Verrips E: Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *Journal of clinical epidemiology* 1998, 51(11):1055-1068.
- (23) Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM: A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med* 1997, 27(2):363-370.
- (24) Bjelland I, Dahl AA, Haug TT, Neckelmann D: The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002, 52(2):69-77.
- (25) Axford J, Butt A, Heron C, Hammond J, Morgan J, Alavi A, Bolton J, Bland M: Prevalence of anxiety and depression in osteoarthritis: use of the Hospital Anxiety and Depression Scale as a screening tool. *Clinical rheumatology* 2010, 29(11):1277-1283.
- (26) Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G: Dimensional assessment of chronic fatigue syndrome. *JPsychosomRes* 1994, 38(5):383-392.
- (27) McHorney CA, Tarlov AR: Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995, 4(4):293-307.
- (28) Streiner DLN, G.R.: Measuring change, Responsiveness and sensitivity to change. In: *Health measurement scales*. 4 edn. Edited by Streiner DLN, G.R. Oxford, UK: Oxford University Press; 2008: 277-288.
- (29) Gabel CP, Melloh M, Burkett B, Michener LA: Lower limb functional index: development and clinimetric properties. *Physical therapy* 2012, 92(1):98-110.
- (30) Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1(8476):307-310.

- (31) Donoghue D, Stokes EK: How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. *J Rehabil Med* 2009, 41(5):343-346.
- (32) Mokkink LB, Terwee CB, Patrick LD, Alonso J, Stratford PW, Knol DL, Bouter LM, De Vet HC: COSMIN-manual. In.; 2010.
- (33) Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC: Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology* 2007, 60(1):34-42.
- (34) Meng X, Rosenthal R, Rubin DB: Comparing Correlated Correlation Coefficients. *Quantitative Methods in Psychology* 1992, 111(1):4.
- (35) Angst F: The new COSMIN guidelines confront traditional concepts of responsiveness. *BMC Med Res Methodol* 2011, 11:152; author reply 152.
- (36) Guyatt GH, Kirshner B, Jaeschke R: Measuring health status: what are the necessary measurement properties? *Journal of clinical epidemiology* 1992, 45(12):1341-1345.
- (37) de Vet HC, Bouter LM, Bezemer PD, Beurskens AJ: Reproducibility and responsiveness of evaluative outcome measures. Theoretical considerations illustrated by an empirical example. *Int J Technol Assess Health Care* 2001, 17(4):479-487.
- (38) Bessette L, Sangha O, Kuntz KM, Keller RB, Lew RA, Fossel AH, Katz JN: Comparative responsiveness of generic versus disease-specific and weighted versus unweighted health status measures in carpal tunnel syndrome. *Med Care*. 1998, 36(4):491-502.
- (39) Stratford PW, Binkley FM, Riddle DL: Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther* 1996, 76(10):1109-1123.
- (40) Schwartz CE, Andresen EM, Nosek MA, Krahn GL: Response shift theory: important implications for measuring quality of life in people with disability. *Arch Phys Med Rehabil* 2007, 88(4):529-536.
- (41) Bombardier C, Melfi CA, Paul J, Green R, Hawker G, Wright J, Coyte P: Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. *Med Care* 1995, 33(4 Suppl):AS131-144.
- (42) Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM: Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord* 2003, 4:10.
- (43) Ponterotto JG, Ruckdeschel DE: An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Percept Mot Skills* 2007, 105(3 Pt 1):997-1014.

3

Joint-Pain Comorbidity, Health Status, and
Medication Use in Hip and Knee
Osteoarthritis: A Cross-Sectional Study.

ABSTRACT

Objective. To determine the prevalence of joint-pain comorbidities in individuals with hip or knee osteoarthritis (OA) and to assess the differences in characteristics of people with and without joint-pain comorbidities.

Methods. In this cross-sectional study, individuals referred to secondary care for treatment of hip/knee OA completed questionnaires to determine socio-demographic characteristics, disease-related outcomes and joint-pain comorbidities. Joint-pain comorbidity was defined as pain perceived in a joint, other than the index joint, for more than half of the days in the preceding month. To compare differences in patient- and disease-related characteristics between participants with and without joint-pain comorbidities we performed analyses of covariance and logistic regression.

Results. A total of 401 individuals, 117 with hip and 284 with knee OA, returned the questionnaire (response rate 82%); mean age 58 (SD=13) and 58% female. Fifty-eight percent of the participants reported symptoms in ≥ 1 other joint. Participants with joint-pain comorbidities were more likely female, less educated and have more medical comorbidities. Individuals with joint-pain comorbidities reported unfavorable outcomes on pain, functioning, fatigue, distress, and health-related quality of life compared with patients without joint-pain comorbidities (all $p < 0.001$). Moreover, use of NSAIDs ($p = 0.038$), opioids ($p = 0.010$), and supplements ($p = 0.019$) was higher in the group with joint-pain comorbidities.

Conclusion. Our results indicate that individuals with joint-pain comorbidities represent a clinically relevant and large subgroup of people with OA of the knee or hip. We recommend addressing joint-pain comorbidities in both research and clinical practice.

Hoogeboom TJ, den Broeder AA, Swierstra BA, de Bie RA, van den Ende CHM.
*Joint-Pain Comorbidity, Health Status, and Medication Use in Hip and Knee Osteoarthritis:
A Cross-Sectional Study.*
Arthritis Care and Research 2012;64:1,54-58

INTRODUCTION

A growing body of evidence demonstrates that in the general population multi-site chronic pain is more common than single-site chronic pain (1, 2). However, as yet the extent and consequences of joint-pain comorbidities in a hip or knee osteoarthritis (OA) population have been addressed in only one study (3). This study shows that joint-pain comorbidities are common in knee OA and that individuals with knee OA and low back pain had less favorable health status compared to individuals without low back pain (3). Some authors suggested that to some extent disappointing therapy results in the management of hip and knee OA can be explained by the fact that other painful joints sites are neglected (4). A better understanding of the relation between joint-pain comorbidities and physical and psychological health outcomes could help identify patients who might benefit from a more holistic treatment approach. Therefore, the aims of this study were: 1. to determine the prevalence of painful joint sites other than the index joint, and 2. to assess differences in characteristics of people with hip or knee OA with and without joint-pain comorbidities.

METHODS

The study was based on cross-sectional data collected from June to October 2009 in the Sint Maartenskliniek Nijmegen (The Netherlands), a specialized hospital for orthopedics, rheumatology and rehabilitation. Men and women (≥ 18 years), received a questionnaire within 14 days after they visited an orthopedic surgeon for a new episode of complaints due to hip or knee OA. Radiographs of the index joint were taken from all patients, as this is standard procedure. Patients who self-reported concurrent rheumatoid arthritis, fibromyalgia or psoriatic arthritis, were excluded. The Institutional Review Board of the University Medical Centre Nijmegen approved the study (ID: 2009/020).

Socio-demographic characteristics were assessed by standardized questions. All posterior-anterior fixed flexion knee and pelvic radiographs (5) were scored for the index joint by a trained and experienced rheumatologist according to the Kellgren and Lawrence grading scale (6). Validated and reliable questionnaires, were used to assess the following disease-related measures: general functioning (range: 0-100) and pain (range: 0-100) (SF-36; lower scores represent unfavorable outcomes) (7), stiffness (range: 0-100), pain (range: 0-100) and functioning (range: 0-100) of the index joint (WOMAC; subscales scores were transformed to a 0-100 scale: $100 - [\text{actual raw score} \times 100 / \text{possible raw score range}]$; lower scores represent unfavorable outcomes), fatigue (range: 8-56, cut-off severely fatigued: ≥ 35) (Checklist Individual Strength (CIS)) (8), self-efficacy (range: 10-40) (General Self-efficacy Scale (GSES); lower scores represent unfavorable outcomes) (9), kinesiophobia

(range: 17-68) (Tampa Scale for Kinesiophobia (TSK); lower scores represent favorable outcomes) (10), anxiety (range: 0-21) and depression (range: 0-21; cut-off depressive symptoms: ≥ 8) (Hospital Anxiety and Depression Scale (HADS); lower scores represent favorable outcomes) (11), and health related quality of life (range: 0-1) (EQ-5D; lower scores represent unfavorable outcomes) (12). Information about health service utilization (HSU) was obtained by asking straightforward questions. (Example of a question: *Have you ever been in contact with a physiotherapist for your knee or hip complaints?*). We assessed contact with a general practitioner, physiotherapist and psychologist. We also assessed analgesic use (i.e. paracetamol, NSAIDS and opioids) and complimentary/alternative care (CAM) use (e.g. acupuncture, healers, dietary supplements, etc).

Finally, the number of painful joint sites (i.e. feet, ankles, knees, hips, lumbar spine, thoracic spine, cervical spine, shoulders, elbows, wrists and hands) was assessed by the following two straightforward questions: “Do you experience symptoms from your (left/right) *joint*?” and “If so, did you experience symptoms of your (left/right) *joint* for more than half of the time during the last month?” A joint was counted as “painful” in case of a positive answer to the second question. Joint-pain comorbidity was defined as pain perceived in a joint, other than the index joint, for more than half of the days in the preceding month.

Table 1. Prevalence of joint-pain comorbidities (pain or aching on more than half the days in the preceding month) by location.*

Pain location	Knee group (n = 284)			Hip group (n = 117)		
	Both sides	Single side	Total	Both sides	Single side	Total
<i>Spine pain</i>						
Lumbar pain	-	66 (23%)	66 (23%)	-	43 (37%)	43 (37%)
Thoracic pain	-	18 (6%)	18 (6%)	-	12 (10%)	12 (10%)
Neck pain	-	49 (17%)	49 (17%)	-	18 (15%)	18 (15%)
<i>Extremity pain</i>						
Shoulder	25 (9%)	22 (8%)	47 (17%)	10 (9%)	7 (7%)	17 (16%)
Elbow	7 (2%)	13 (5%)	20 (7%)	2 (2%)	3 (3%)	5 (5%)
Wrist	14 (5%)	12 (4%)	26 (9%)	4 (4%)	5 (5%)	9 (9%)
Hand	23 (8%)	16 (6%)	39 (14%)	6 (6%)	4 (4%)	10 (10%)
Hip	15 (5%)	26 (9%)	41 (14%)	31 (26%)	64 (55%)	95 (81%)
Knee	81 (29%)	160 (56%)	241 (85%)	10 (9%)	26 (24%)	36 (33%)
Ankle	9 (3%)	22 (8%)	31 (11%)	5 (5%)	7 (7%)	12 (12%)
Foot	18 (6%)	21 (7%)	39 (14%)	7 (7%)	6 (6%)	13 (13%)

* Values are the number (percentage) of participants.

A total of 384 participants were needed to sufficiently power the study; assuming a 50% prevalence rate of joint-pain comorbidity, an error margin of 5% and 95%-confidence

limits. Descriptive statistics were used to describe the amount of missing data and the study population. Frequencies and percentages were used to report the prevalence of joint-pain comorbidities. On basis of complete case data, socio-demographic data between the group with and without joint-pain comorbidities were compared using t-test and chi-square statistics. To compare disease-related characteristics of participants with and without joint-pain comorbidities we performed analyses of covariance and logistic regression when appropriate, adjusting for age, sex, education, medical comorbidity and index joint (i.e. hip or knee). Sub-group analyses were performed for the isolated hip and knee groups.

RESULTS

A total of 401 people completed and returned the questionnaire (response rate 82%). The non-responder group (of whom 25% were contacted by telephone) did not differ significantly from the responder group on demographic variables. Three hundred forty four (86%) participants had less than three missing data points. The study sample was 58% female, with a mean \pm SD age of 58 ± 13 years and a mean \pm SD BMI of 27 ± 5 kg/m². A total of 284 participants were referred for knee OA and 117 for hip OA; 73 (18%) participants reported both hip and knee complaints. Seventy-six percent of the participants had a Kellgren & Lawrence grade of 2 or higher.

Table 2. Patient- and disease-characteristics of individuals with hip and/or knee OA with and without joint-pain comorbidities.*

	No JPC (n=170)	≥ 1 JPC (n=231)	Mean diff / Odds Ratio (95%-CI)	<i>p</i>
Demographics				
Age, mean \pm SD yrs	57.7 \pm 13.7	58.4 \pm 12.2	-0.72 (-3.27 ; 1.83)	0.579
Sex, women	87 (51%)	144 (62%)		0.025
BMI, mean \pm SD kg/m ²	27.1 \pm 4.5	27.3 \pm 4.9	0.34 (-0.65 ; 1.34)	0.496
Medical comorbidity [¶]				
Diabetes mellitus	9 (5%)	19 (8%)	0.83 (0.34 ; 2.00)	0.674
Cardiovascular disease	13 (8%)	34 (15%)	0.45 (0.21 ; 0.96)	0.040
Respiratory disease	6 (4%)	21 (9%)	0.31 (0.11 ; 0.87)	0.026
Kidney/Liver disease	2 (1%)	2 (1%)	0.71 (0.10 ; 5.12)	0.732
Gastrointestinal complaints	8 (5%)	24 (11%)	0.39 (0.16 ; 0.94)	0.036
Education, secondary school	155 (95%)	191 (87%)		0.036
Additionally insured	146 (88%)	202 (92%)		0.207
Income, >€40,000 per year	55 (36%)	55 (29%)		0.118
Currently employed	87 (53%)	94 (43%)		0.062

Abbreviations: JPC, joint-pain comorbidity; OA, osteoarthritis; * Values are the number (percentage) of participants unless indicated otherwise. [¶] Adjusted for age, sex, index joint, and education.

Fifty-eight percent (n=231) of the participants reported symptoms in more than one joint for more than half of the time during the previous month (Table 1). The median (interquartile range) number of affected joints was 2 (1-4) joints. Peripheral regions most and least often affected in our total sample, other than hip (34%) and knee (69%), were the shoulders (16%) and elbow joints (6%), respectively. The lumbar spine (27%) was most often reported as painful when considering the axial joints. The number of patients with knee OA reporting comorbid pain hip pain was significantly lower number than the number of individuals with hip OA reporting comorbid knee pain (14% vs. 33%, respectively). Also, pain in the lumbar spine region was more frequent in the patients referred for hip OA than those referred for knee OA (23% vs. 37%, respectively).

Participants with painful joints other than the index joints were more likely to be female, less educated and have more medical comorbidities (Table 2). Individuals with joint-pain elsewhere showed significantly more depressive symptoms (HADS depression score ≥ 8) (35% vs. 12%, $p < 0.001$) and were more severely fatigued (CIS fatigue score ≥ 35) (56% vs. 31%, $p < 0.001$) than those without joint-pain comorbidities; odds ratios (95%-CI) 3.56 (2.00 ; 6.35) and 2.34 (1.49 ; 3.70) resp. (Table 3). Moreover, pain (SF-36 & WOMAC), symptoms (WOMAC), functioning (SF-36 & WOMAC), quality of life (EuroQoL) and the value of own health (EuroQoL) were significantly worse in individuals with joint-pain comorbidities. No statistically significant differences between groups were found in radiographic severity (0.791), kinesiophobia ($p = 0.727$), duration of complaints ($p = 0.089$) and self-efficacy ($p = 0.099$).

NSAIDS (54% vs. 64%, $p = 0.038$), opioids (11% vs. 24%; $p = 0.010$) and supplements use (29% vs. 40%, $p = 0.019$) were significantly higher in the group that reported joint-pain comorbidities than those who did not (Table 3); odds ratios (95%-CI) 1.60 (1.03 ; 2.49), 2.21 (1.21 ; 4.05) and 1.75 (1.10 ; 2.79) respectively. No statistically significant differences were found in the number of patients that contacted a physiotherapist (75% vs. 79%, $p = 0.397$) or general practitioner (93% vs. 92%, $p = 0.618$) for their symptoms.

We repeated all analyses for the hip (n=117) and knee group (n=284) separately. Analyses in the knee group yielded results similar to the primary analyses. For the hip group, we did find some differences from the primary analyses. Compared with the primary analyses, no significant difference between hip OA patients with and without joint-pain comorbidities were found in use of NSAIDs, opioids and supplements, pain (SF-36), functioning (WOMAC & SF-36) and quality of life (EQ-5D). However, although not significant, the directions of the differences were similar to the primary analyses, except for NSAIDs use.

Table 3. Disease characteristics of individuals with hip and/or knee OA with and without JPC.*

	No JPC (n=170)	≥1 JPC (n=231)	Mean diff (95%-CI)	Odds Ratio (95%-CI)
Health and disease related status[¶]				
Fatigued, CIS score ≥35, n (%)	51 (31%)	126 (56%)	-	2.34 (1.49 ; 3.70)
Depressive symp, HADS ≥8, n (%)	20 (12%)	79 (35%)	-	3.56 (2.00 ; 6.35)
Anxiety ^{S,‡}	4.3 ± 3.4	5.9 ± 4.1	1.44 (0.64 ; 2.23)	-
Kinesiophobia [‡]	38.6 ± 7.8	39.3 ± 8.1	0.30 (-1.37 ; 1.96)	-
Self-efficacy [†]	32.8 ± 4.9	31.4 ± 5.6	-0.91 (-2.00 ; 0.17)	-
Quality of life ^{¥,†}	0.71 ± 0.17	0.63 ± 0.19	-0.06 (-0.10 ; -0.02)	-
Value own health ^{¥,†}	6.1 ± 1.7	5.8 ± 1.6	-0.64 (-0.99 ; -0.30)	-
SF-36				
Bodily pain [†]	45.8 ± 20.5	35.8 ± 18.4	-8.80 (-12.85 ; -4.75)	-
Physical functioning [†]	48.2 ± 23.1	36.9 ± 20.4	-9.17 (-13.60 ; -4.74)	-
OA and symptoms[¶]				
Complaint period, ≥5 yrs, n (%)	88 (55%)	139 (66%)	-	1.47 (0.94 ; 2.29)
K&L, ≥2, n (%)	129 (76%)	174 (76%)	-	0.93 (0.57 ; 1.54)
WOMAC [€]				
Pain [†]	61.0 ± 21.6	49.7 ± 20.0	-9.82 (-14.17 ; -5.48)	-
Stiffness [†]	52.9 ± 21.5	40.6 ± 21.1	-11.73 (-16.19 ; -7.26)	-
Function [†]	60.2 ± 22.0	48.8 ± 20.3	-9.75 (-14.29 ; -5.21)	-
HCU[¶], n (%)				
GP contact (≥1)	157 (93%)	209 (92%)	-	0.81 (0.37 ; 1.78)
PT contact (≥1)	126 (75%)	180 (79%)	-	1.26 (0.75 ; 2.12)
CAM				
Contact	19 (11%)	42 (18%)	-	1.57 (0.84 ; 2.95)
Supplement use	50 (29%)	93 (40%)	-	1.65 (1.04 ; 2.63)
Acetaminophen use	98 (58%)	155 (68%)	-	1.32 (0.85 ; 2.06)
NSAID use	92 (54%)	146 (64%)	-	1.60 (1.03 ; 2.49)
Opioids use	18 (11%)	55 (24%)	-	2.29 (1.25 ; 4.16)

Abbreviations: BMI, Body Mass Index; CAM, Complementary and alternative medicine; CIS, Checklist Individual Strength; HADS, Hospital Anxiety and Depression Scale; HCU, Health care utilization; JPC, joint-pain comorbidity; OA, osteoarthritis; PT, Physiotherapy; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index. * Values are mean ± standard deviation of participants unless indicated otherwise. [¶] Adjusted for age, sex, index joint, education and medical comorbidity. [¥] EQ-5D questionnaire. ^S HADS questionnaire. [†] Lower scores represent unfavorable outcomes. [‡] Higher scores represent unfavorable outcomes. [€] WOMAC subscales are transformed and range from 0-100.

DISCUSSION

In this study we showed that the majority of people referred for hip or knee OA experience joint pain elsewhere. The physical and psychological burden of symptoms related to knee or hip OA is significantly higher in individuals with painful joint sites elsewhere than individuals who did not report joint-pain comorbidity, especially in the group of patients with knee OA. Moreover, NSAIDS, opioids and supplements use were also considerably higher in the group with joint-pain comorbidities, whereas no differences in general practitioner and physiotherapy contacts were observed. Thus, we identified a relevant subgroup of patients possibly requiring more intensive and complex treatment interventions.

In contrast to previous studies (3, 13), the prevalence of joint-pain comorbidities in our study (58%) was lower. Dissimilarities in prevalence might be explained by differences in the method of determining joint-pain comorbidities. We defined a joint site as painful when a participant reported pain for more than half of the days in the preceding month whereas Croft *et al* defined a joint site as painful in case of pain in one or more day in the last month (13). Differences in the prevalence of painful joint sites between our and Suri *et al*'s study, could be explained by the fact they use a manikin to determine joint-pain comorbidities. Although both methods (written questions and manikin) give valid results, the use of manikin is known to result in higher prevalence of musculoskeletal pain (14).

Individuals with hip or knee OA with joint-pain comorbidities are more fatigued, experience lower quality of life, more bodily pain, deprived functioning and might even be at risk for depressive symptoms compared with people without joint-pain comorbidities. The magnitude of these differences is substantial and can be considered as clinically relevant. For instance, the absolute number of participants who had a fatigue score above 35 points (8) - indicating severe levels of fatigue - was 23% higher in participants with than participants without joint-pain comorbidities. Moreover, patients with pain elsewhere use more and stronger analgesics than patients without joint-pain comorbidities and are more likely to use dietary supplements of controversial efficacy. In conclusion, pain elsewhere in hip or knee OA is associated with compromised health status and more (unfavorable) forms of medication and supplement use. We therefore concur with the recommendation that in both research and clinical practice joint-pain comorbidities should be accounted for in hip and knee OA (4).

In our opinion differences in health status and quality of life between individuals with and without joint-pain comorbidities cannot be attributed to differences in the progression of OA as disease duration and Kellgren and Lawrence grading were comparable in both groups. Our results do not allow interpretation of the causal relationships between joint-pain comorbidities and health status, due to the cross-sectional nature of the study.

Our data showed that the proportion of patients who contacted health care providers was similar between the two groups whereas stronger medication use and dietary supplement intake was more prevalent in the group with joint-pain comorbidities. Considering the substantially higher burden of disease in multiple domains in the subgroup of patients with joint-pain comorbidities (e.g. more medical comorbidities, depressed mood, anxiety and severe fatigue), it is conceivable that these patients may require, next to medication, more complex and intensive treatment modalities than patients without joint-pain comorbidities. Our results might be a starting point to tailor treatment for this category of patients.

A possible limitation of our study is that we did not collect data on the nature of joint-pain comorbidities other than the index joint. Thus, it is unknown to what extent those complaints are related to OA (e.g. generalized OA) or possibly to other diagnoses. Further research should focus on the cause and nature of joint-pain comorbidities in knee or hip OA. Another limitation of our study is that we recruited individuals from one centre, a large specialized hospital in the Netherlands which might hamper the generalizability of our findings. However, as patient and disease characteristics of our sample resemble those of the Osteoarthritis Initiative Incidence subcohort (15) we believe that our results are generalizable to patients with hip or knee OA referred to secondary care.

Another limitation of our study is that we did not power the study specifically for subgroup analyses. Due to lack of power it remains unclear whether the differences in results between the hip and knee subgroups reflect true differences. Nevertheless, considering that we are the first to investigate joint-pain comorbidities in hip OA, we think that additional research is warranted.

In conclusion, in this study we generated empirical data confirming the expectation that people with joint-pain comorbidities represent a clinically relevant and relatively large subgroup. In this subgroup of patients the burden of disease is considerably higher, medical comorbidities are more prevalent and medication use is significantly higher. Because our study does not allow inferences about the nature and causality of joint-pain comorbidities in OA further research is warranted.

REFERENCE LIST

- (1) Carnes D, Parsons S, Ashby D, Breen A, Foster NE, Pincus T, et al. Chronic musculoskeletal pain rarely presents in a single body site: results from a UK population study. *Rheumatology(Oxford)* 2007;46(7):1168-70.
- (2) Kamaleri Y, Natvig B, Ihlebaek CM, Bruusgaard D. Localized or widespread musculoskeletal pain: does it matter? *Pain* 2008;138(1):41-6.
- (3) Suri P, Morgenroth DC, Kwok CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. *Arthritis Care Res(Hoboken)* 2010;62(12):1715-23.
- (4) Conaghan PG, Birrell F, Burke M, Cumming J, Dickson J, Dieppe P, et al. Osteoarthritis: National clinical guideline for care and management in adults. 2008.
- (5) Peterfy C, Li J, Zaim S, Duryea J, Lynch J, Miaux Y, et al. Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test-retest reproducibility. *Skeletal Radiol* 2003;32(3):128-32.
- (6) Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthritis. *AnnRheumDis* 1957;16(4):494-502.
- (7) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *MedCare* 1992;30(6):473-83.
- (8) Beurskens AJ, Bultmann U, Kant I, Vercoulen JH, Bleijenberg G, Swaen GM. Fatigue among working people: validity of a questionnaire measure. *OccupEnvironMed* 2000;57(5):353-7.
- (9) Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. *JPsychol* 2005;139(5):439-57.
- (10) Houben RM, Leeuw M, Vlaeyen JW, Goubert L, Picavet HS. Fear of movement/injury in the general population: factor structure and psychometric properties of an adapted version of the Tampa Scale for Kinesiophobia. *JBehavMed* 2005;28(5):415-24.
- (11) Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta PsychiatrScand* 1983;67(6):361-70.
- (12) Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. *BrJRheumatol* 1997;36(7):786-93.
- (13) Croft P, Jordan K, Jinks C. "Pain elsewhere" and the impact of knee pain in older people. *Arthritis Rheum* 2005;52(8):2350-4.
- (14) van den Hoven LH, Gorter KJ, Picavet HS. Measuring musculoskeletal pain by questionnaires: the manikin versus written questions. *EurJPain* 2010;14(3):335-8.
- (15) Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. *AnnRheumDis* 2009;68(9):1413-9.

4

Longitudinal study of the impact of joint-pain
comorbidities on quality of life and activity
levels in people with knee osteoarthritis: Data
from the Osteoarthritis Initiative.

ABSTRACT

Objectives. Joint-pain comorbidity (JPC) is common in individuals with knee osteoarthritis (OA). This study investigates the longitudinal association between JPC and health-related quality of life (HRQoL) and physical activity levels in individuals with knee OA.

Methods. Data from the progression cohort of the Osteoarthritis Initiative (n=1233; age=61 and 58% female) were analyzed. JPC was considered present if individuals reported pain in ≥ 3 joint groups, including the knee joints. Health-related quality of life (HRQoL) was assessed using the Knee Injury and Osteoarthritis Outcome Score (KOOS) subscale Quality of Life and self-reported physical activity was determined using the Physical Activity Scale for the Elderly (PASE). GEE-analyses were performed, adjusted for age, sex, duration of complaints, medical comorbidity, and physical and mental functioning.

Results. Over the four-year period, 32% participants never reported JPC, whereas 12% always reported JPC. GEE modeling demonstrated that having JPC was negatively associated with HRQoL; regression coefficient β (95%-Confidence Interval) -3.57 (-4.69 ; -2.44) and not associated with physical activity; -1.32 (-6.61 ; 3.98).

Conclusion. Considering the impact of JPC on HRQoL of individuals with knee OA, the assessment of JPC in individuals with knee OA might be daily routine.

Hoogeboom TJ, den Broeder AA, de Bie RA, van den Ende CHM.

Longitudinal study of the impact of joint-pain comorbidities on quality of life and activity levels in people with knee osteoarthritis: Data from the Osteoarthritis Initiative.

Rheumatology (Oxford), 2012.

INTRODUCTION

Chronic widespread pain is common in the general population (ranging between 11-24%) [1]. In persons with knee osteoarthritis (OA) joint-pain comorbidity (JPC) is considered even more common (prevalence over 50%) [2-5] and burdensome [4-5]. Recent studies demonstrated that JPC was associated with lower levels of physical and psychological health [4-5] than people without JPC. However, due to the cross-sectional nature of these data, it remains unclear whether the association between health related quality of life (HRQoL) and JPC is longitudinally consistent over time. If this relation is indeed apparent, it could be valuable to advise clinicians to address JPC to optimize treatment outcomes in this group of patients.

Hoozeboom *et al* (2012) demonstrated that individuals reporting JPC had twice the odds of having other medical comorbidities, such as cardiovascular and respiratory diseases [5]. One of the most important therapy strategies to prevent these comorbid diseases is the promotion of physical activity. We hypothesized that individuals with JPC, because of the widespread pain, might be specifically prone for physical inactivity [6-8].

Therefore, the first aim of this study is to confirm the longitudinal association between JPC and HRQoL and the second aim is to study the longitudinal association between JPC and physical activity levels in individuals with established knee OA.

PATIENTS AND METHODS

The Osteoarthritis Initiative (OAI) is a publicly and privately funded prospective four-year longitudinal cohort study (available for public access at <http://www.oai.ucsf.edu>) [9]. We included data from the progression sub-cohort. The progression sub-cohort comprises 1,390 persons with symptomatic knee OA in one or both knees. Exclusion criteria for the OAI population included self-reported rheumatoid arthritis, SLE, psoriatic arthritis, ankylosing spondylitis or another inflammatory arthritis (defined as self-report of a physician diagnosis and ever use of specific prescription medications), MRI contraindication, (plans for) bilateral total knee joint replacement, and co-morbid conditions that could interfere with the ability to participate in a 4-year study [9]. To ensure a homogeneous OA cohort, we furthermore excluded participants who self-reported other forms of inflammatory arthritis (n=157); resulting in a total of 1233 cases.

Baseline data on age, sex, marital state, number of medical comorbidities and duration of the disease were acquired. Data on JPC, health-related quality of life and physical activity were acquired on five yearly time points; at baseline, 1-year, 2-year, 3-year and 4-year follow-up; OAI-databases v0.2.2, v1.2.1, v3.2.1, v5.2.1, and v6.2.1 respectively. Joint-pain

was considered present when a participant reported pain, aching or stiffness in a joint (i.e. neck, thoracic spine, lower back, shoulder, elbow, wrist, hand, hip, knee, ankle and foot) for more than half of the days in the past 30 days [4], by use of the following question: *‘During the past 30 days, which of these joints have had pain, aching, or stiffness on most days?’*. Self-reported HRQoL was measured using the subscale Quality of Life of the Knee Osteoarthritis Outcome Score (KOOS) [10]. Self-reported activity levels were measured by use of the Physical Activity Scale for the Elderly (PASE) [11].

Participants were stratified into two categories, having JPC (at least 3 affected joint groups) or having no JPC (2 or fewer affected joint groups). This classification was used to establish for a clear threshold; one often used to define wide-spread pain or generalized osteoarthritis [1, 12]. Consequently participants were pragmatically categorized into the following four groups: 1. ‘never JPC’, 2. ‘sometimes JPC’ (JPC on one or two time points), 3. ‘often JPC’ (JPC on three or four time points), and 4. ‘always JPC’ (JPC on all five time points); to allow us to study the impact of persistence of JPC and to study a possible dose-response relationship. Descriptive statistics were used to describe 1) the different groups, 2) the drop-out rate, and 3) the number of missing values. Missing data mechanisms were studied by use of indicator variables [13]. As the data were considered at least missing at random, missing data were imputed by use of multiple imputation by chained equations, to increase power, enable more efficient analyses and reduce bias [14]. All analyses were performed on 10 multiple imputed datasets [15] and combined using Rubin’s Rules [16]. HRQoL and physical activity data were plotted over time for each of the four groups (mean \pm 95%-Confidence Interval (95%-CI)). By use of generalized estimating equations (GEE) modeling, we studied the longitudinal association between the presence of JPC (yes/no) and HRQoL / physical activity. To do so, we first built a base model comprising JPC, HRQoL / physical activity and time. Consequently, we built controlled models, i.e. base models adjusted for age, sex, number of medical co-morbidities, duration of complaints, and knee pain (WOMAC subscale pain) [13]. Quasi-likelihood under the independence model criterion was used to find an acceptable working correlation structure for the models [17]. Sensitivity analyses were performed on the complete case data (i.e. non-imputed data set). All statistical analyses were carried out using statistical package Stata/IC 12.

RESULTS

Of all 1233 participants (mean age \pm SD: 61 \pm 9 years and 58% female) 44% of the participants reported 3 or more painful joint groups at baseline. At one year 99 (8%) people were lost to follow-up, after year two, three and four years this was respectively 148 (12%), 170 (14%) and 174 (14%). The number of missing values for KOOS subscale quality of life was negligible (<1%) and ranged from 3% to 8% for the PASE variable. In table 1 baseline

characteristics of the four groups are presented. Over the four year period, 32% participants never reported JPC, whereas 12% reported JPC at each of the measurement points.

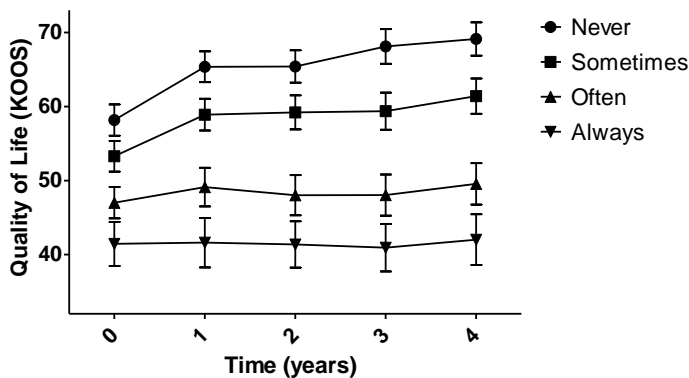
Table 1. Baseline characteristics of the four groups of individuals with or without JPC.*

	Never group	Sometimes group	Often group	Always group	<i>p</i> -value
N (%)	315 (32)	331 (33)	234 (23)	119 (12)	
Age, yrs	62 (9)	62 (9)	62 (9)	60 (9)	0.39
Sex, ♀%	43%	55%	64%	71%	<0.01
Complaints duration >5y, %	38%	48%	52%	64%	<0.01
Medical co-morbidities > 1, %	4%	7%	10%	10%	<0.01
WOMAC stiffness	2.38 (1.6)	2.88 (1.7)	3.42 (1.6)	3.49 (1.5)	<0.01
WOMAC pain	4.03 (3.2)	5.21 (3.7)	6.57 (3.8)	7.24 (3.9)	<0.01
WOMAC activities	12.4 (10.8)	16.8 (12.1)	20.1 (12.2)	23.6 (12.6)	<0.01

Abbreviations: JPC = joint pain comorbidity, SD = standard deviation, WOMAC = Western Ontario and McMaster's University Osteoarthritis Index. *Values are mean (SD) unless indicated otherwise.

In figure 1 we have graphed the relationship between the persistence of JPC and HRQoL. A dose-response relation is visible, namely the participants in the 'never JPC' group reported better HRQoL scores than the 'sometimes JPC' group who reported better HRQoL scores than the 'often JPC' group. Individuals who reported JPC at each of the measurement moments had the worst HRQoL scores (Figure 1).

Figure 1. HRQoL scores plotted over time for the four JPC groups.



The base model demonstrated that the presence of JPC was statistically significantly negatively associated with HRQoL, regression coefficient β (95%-CI) -14.50 (-16.14 ; -12.87), indicating individuals with JPC had a lower score of 14.5 points on the KOOS HRQoL than those without JPC. The corrected model demonstrated orientated point

estimates for having JPC and the influence of time JPC similar to the base model, respectively -3.57 (-4.69 ; -2.44). Time was a positive factor for HRQoL in both models, indicating an increase in HRQoL over time; base model 1.25 (0.94 ; 1.55) and corrected model 0.79 (0.54 ; 1.03). Complete case analysis produced a similar plot with broader and overlapping confidence intervals. In addition, GEE on complete case data produced similar statistically significant associations with smaller β -values.

Physical activity decreased over time, with similar slopes for each of the groups and confidence intervals overlapping, with the exception of the ‘never’ and ‘always’ groups. GEE analyses revealed in the base model a negative, longitudinal association between physical activity levels and having JPC; -9.88 (-16.00 ; -3.76). However, the corrected model demonstrated that having JPC was not associated with physical activity levels (-1.32 (-6.61 ; 3.98)). In both the base and corrected model time was a negative factor, -4.64 (-5.85 ; -3.44) and -4.79 (-6.01 ; -3.57). Visual and statistical analyses on complete case data yielded similar results as the multiple imputed data.

DISCUSSION

Our results confirm that JPC is highly prevalent in individuals with knee OA. Moreover, in line with our hypothesis we found that JPC was negatively associated with HRQoL, whereas our hypothesis that JPC was associated with physical activity could not be confirmed. The unfavorable association between health-related outcomes and JPC in individuals with OA has already been established cross-sectionally [4-5]. However, we are the first to demonstrate this relation longitudinally, thus providing a more precise estimate of the strength of this relationship.

Our findings can have some implications for clinical care in knee OA. We propose that the assessment of JPC should be part of routine (clinical and research) practice, and that health care providers should be aware of the association between JPC and HRQoL when treating individuals with knee OA. In case JPC is present, a clinician or researcher should be aware that this person could be clinically and prognostically different from a patient without JPC. Perhaps therapy goals need to be adjusted accordingly, as focusing merely on the most affected joint might result in disappointing therapy results. The latter needs to be confirmed in intervention studies [18]. Unfortunately, research on the management of JPC in individuals with OA is still scarce [19] and in need of further study [12].

We are the first to study the association between having JPC and physical activity. Our results indicate that physical activity decreases markedly over a four-year period (approximately 11 per cent ($p < 0.01$) for the whole group (DNS)). However, we could not confirm our hypothesis that having JPC contributed to an extra decline in physical activity while accounting for socio-demographics, medical comorbidities, and knee pain. On the

other hand, it should be noted that the validity of the PASE questionnaire is questioned [20]; necessitating further study on JPC in OA and objectively measured physical activity. Although the strengths of this study include the large, representative sample of individuals from a variety of cultural backgrounds, this study also has some limitations. Patient-reported outcomes are prone for a phenomenon called response shift; the potential of subjects' views, values, or expectations to change over the course of a study, thereby adding an additional factor of change on results [21]. Response shift might explain why HRQoL improved over time in our analyses. Another limitation is the lack of insight on the nature of the joint pain, as only global questions were asked about the chronicity of JPC while acute joint pain was not collected.

In conclusion, JPC has a significant impact on the HRQoL of individuals with knee OA. Therefore, both researchers and clinicians should assess JPC in their daily practice when working with individuals with knee OA. In addition, the general trend towards physical inactivity is reason for concern.

ACKNOWLEDGEMENTS

The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This manuscript was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners.

REFERENCE LIST

- (1) Cimmino MA, Ferrone C, Cutolo M. Epidemiology of chronic musculoskeletal pain. *Best Pract Res Clin Rheumatol* 2011;25(2):173-83.
- (2) Forestier R, Francon A, Briole V, Genty C, Chevalier X, Richette P. Prevalence of generalized osteoarthritis in a population with knee osteoarthritis. *Joint Bone Spine* 2011;78(3):275-8.
- (3) Gunther KP, Sturmer T, Sauerland S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. *Ann Rheum Dis* 1998;57(12):717-23.
- (4) Suri P, Morgenroth DC, Kwok CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. *Arthritis Care Res.(Hoboken.)* 2010;62(12):1715-23.
- (5) Hoogbeem TJ, den Broeder AA, Swierstra BA, de Bie RA, van den Ende CH. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study. *Arthritis Care Res (Hoboken)* 2012;64(1):54-8.
- (6) From the Centers for Disease Control and Prevention. Prevalence and impact of chronic joint symptoms--seven states, 1996. *JAMA* 1998;279(24):1940-1.
- (7) Feinglass J, Nelson C, Lawther T, Chang RW. Chronic joint symptoms and prior arthritis diagnosis in community surveys: implications for arthritis prevalence estimates. *Public Health Rep* 2003;118(3):230-9.
- (8) Busija L, Buchbinder R, Osborne RH. Quantifying the impact of transient joint symptoms, chronic joint symptoms, and arthritis: a population-based approach. *Arthritis Rheum* 2009;61(10):1312-21.
- (9) Nevitt M, Felson D, Lester G. The Osteoarthritis Initiative: Protocol for the Cohort Study. In; 2006.
- (10) Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes* 2003;1:17.
- (11) Washburn RA, Ficker JL. Physical Activity Scale for the Elderly (PASE): the relationship with activity measured by a portable accelerometer. *J Sports Med Phys Fitness* 1999;39(4):336-40.
- (12) Hoogbeem TJ, Stukstette MJ, de Bie RA, Cornelissen J, den Broeder AA, van den Ende CH. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. *BMC Musculoskelet Disord* 2010;11:142.
- (13) Twisk JWR. *Applied Longitudinal Data Analysis for Epidemiology: A practical guide*. New York: Cambridge University Press; 2003.
- (14) Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
- (15) Royston P. Multiple imputation of missing values: Update of ice. *Stata Journal* 2005;5(4):527-36.

- (16) Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC.Med.Res.Methodol.* 2009;9:57.
- (17) Cui J. QIC program and model selection in GEE analyses. *The Stata Journal* 2007;7(2):12.
- (18) Miles CL, Pincus T, Carnes D, et al. Can we identify how programmes aimed at promoting self-management in musculoskeletal pain work and who benefits? A systematic review of sub-group analysis within RCTs. *Eur J Pain* 2011;15(8):775 e1-11.
- (19) Conaghan PG, Birrell F, Burke M, et al. Osteoarthritis: National clinical guideline for care and management in adults. In; 2008.
- (20) Svege I, Kolle E, Risberg MA. Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis. *BMC Musculoskelet Disord* 2012;13:26.
- (21) Robling M, Hood K. Response shift, responsiveness or recall bias? *Br J Gen Pract* 2002;52(480):585.

5

The impact of waiting for total joint replacement on pain and functional status:
a systematic review.

ABSTRACT

Objective. To systematically describe changes in pain and functioning in patients with osteoarthritis (OA) awaiting total joint replacement (TJR), and to assess determinants of this change.

Methods. MEDLINE®, EMBASE, CINAHL® and Cochrane Database were searched through June 2008. The reference lists of eligible publications were reviewed. Studies that monitored pain and functioning in patients with hip or knee OA during the waiting list for TJR were analyzed. Data was collected with a pre-specified collection tool. Methodological quality was assessed and a best evidence analysis was performed to summarize results.

Results. Fifteen studies, of which two were of high quality, were included and involved 788 hip and 858 knee patients (mean age 59-72 and mean wait 42-399 days). There was strong evidence that pain (in hip and knee OA) and self-reported functioning (in hip OA) do not deteriorate during a <180 days wait. Conflicting evidence was established for the change on self-reported functioning in patients with knee OA waiting <180 days. Moreover, strong evidence was found for an association between the female gender and intensified pain.

Conclusion. Patients with OA do not experience deterioration in pain or self-reported functional status whilst waiting <180 days for TJR. Changes over a longer waiting period are unclear. To strengthen and complement the present evidence, further high-quality studies are needed, in which preferably also performance-based measures are used.

Hoogbeem TJ, van den Ende CHM, van der Sluis G, Elings J, Dronkers JJ,
Aiken AB, van Meeteren NLU.

*The impact of waiting for total joint replacement on pain and functional status:
a systematic review.*

Osteoarthritis and Cartilage 2009;17,1420-1427.

INTRODUCTION

Annually, over three-quarter of a million surgical total hip (THR) and knee replacements (TKR) are executed in the United States. Due to aging of the population and current Western lifestyle patterns ^{1,2} this number is expected to grow substantially over the next two decades ³. Patients with osteoarthritis (OA) scheduled for an elective major total joint replacement (TJR) already endured a lasting period of gradual functional decline ^{4,5}. Since preoperative functioning is the strongest determinant of postoperative pain and functioning ⁶⁻⁸, it is important to understand whether prolonged wait times are associated with deterioration in the preoperative status. The latter is in accordance with the recommendations of Lane (2007) and Wylde *et al* (2007) who emphasize the need for a better understanding of the epidemiology and progression of OA to facilitate identification of high-risk patients who might benefit from interventions to (a) alter the course of the disease ⁹ or (b) minimize the risk of having a poor outcome in arthroplastic surgery ¹⁰. So far, the evidence on the natural course of pain and functional status in patients scheduled for arthroplastic surgery is not thoroughly summarized. In 2005 Noseworthy *et al* reviewed the relationship between waiting time and the change in health status in patients with OA ¹¹. They concluded that there is some evidence indicating that functional status deteriorates in patients waiting more than six months for joint replacement surgery. This conclusion, however was based upon four studies with very diverse quality. Therefore, the objective of this systematic review of the literature was 1) to describe the change of pain and functional status in patients with OA awaiting THR or TKR, and 2) to assess determinants of this change.

Methods

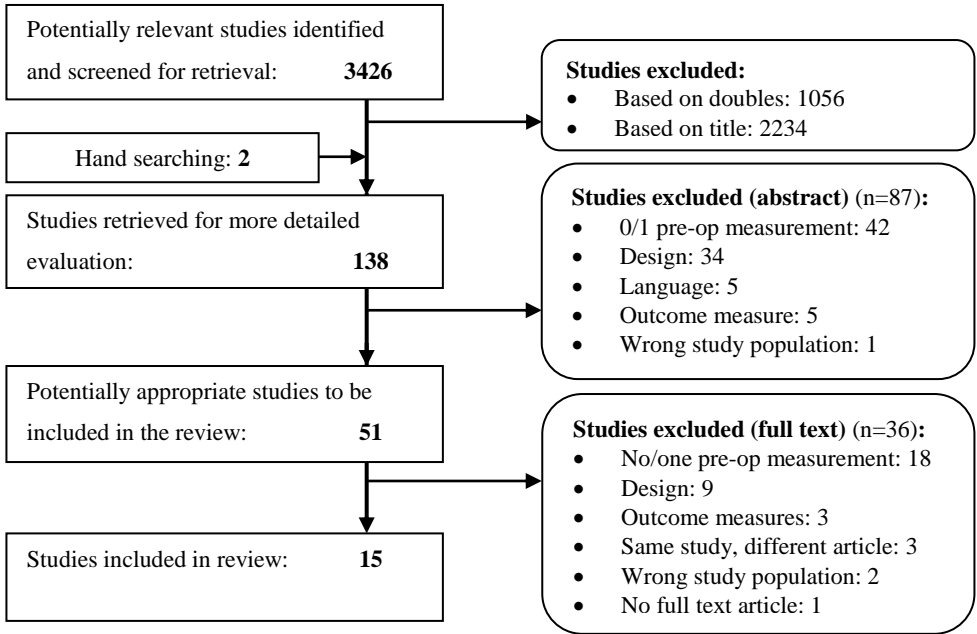
Identification of studies

Initially the online databases MEDLINE®, EMBASE, CINAHL® and the Cochrane Database of Systematic Reviews (through June 2008) were searched. With the Boolean terms ‘AND’ and ‘OR’ we combined the following three groups of key words and MeSH-terms: (1) “Joint Replacement”, “Joint Arthroplasty”, “Arthroplasty, Replacement”, “Hip Replacement”, “Arthroplasty, Replacement, Hip”, “Total Hip”, “Hip Arthroplasty”, “Knee Replacement”, “Arthroplasty, Replacement, Knee”, “Total Knee”, “Knee Arthroplasty”, or “End-stage Osteoarthritis”, (2) “Wait*”, “Queu*”, “Preoperative”, “Preoperative Care”, “Preadmission”, or “Presurgery”, and (3) “Function*” or “Pain” or “Discomfort”. Furthermore to retrieve any relevant articles that the search string may have missed, the

first author hand searched the references of all relevant articles and checked the work of the first and last authors of all eligible articles.

By using the prespecified eligibility criteria, two researchers (TH & CE) screened all electronic citations to select reports for full-text review. Then, TH and GS assessed each full-text article for inclusion. The team reviewed all undecided cases in plenary sessions. Reasons for exclusion of ineligible studies are recorded in the flow diagram (Fig. 1) ¹².

Figure 1. QUOROM flow chart showing the number of studies screened and included in the review.



Selection criteria

Studies were eligible for inclusion in the review if: 1) the data were prospective in nature (i.e. descriptive, cohort, randomized trials, and randomized intervention trials with no intervention in the control group ¹³); 2) changes in pain or functioning (activities and participation, described in this study as functional status) were monitored during the waiting list period before elective THR or TKR (at least two preoperative measuring time points are needed to assess change); 3) the study population consisted of at least 80% of the patients with OA; and 4) the study provided point estimates and measures of variability. We defined the waiting list period as the time between the date of the decision to treat surgically and the actual date of surgery ¹⁴. Studies in languages other than Dutch or English were excluded.

Assessment of methodological quality

A checklist, developed by Tooth *et al* (2005) for the assessment of methodological quality of studies, was used to rate the included studies ¹⁵. This checklist consists of 33 criteria, covering the study rationale, population, recruitment, measurement, biases, data analysis, and generalizability of the results. All 33 criteria were scored on a dichotomous scale ('Yes' or 'No'). Each article was independently rated with the aforementioned checklist by two authors (TH and GS). The interrater reliability was examined with Cohen's Kappa. Disagreements were resolved by a consensus meeting.

Data collection

Study characteristics extracted from eligible papers were: authors, year of publication, study design, setting, methods of data collection, inclusion and exclusion criteria, sample size, description of measurements, waiting list period, outcome measures, prognostic factors and conclusions. These data were extracted using a pre-specified extraction form.

Data analysis

Data analysis methodology was adopted from the recommendations of Van Dijk *et al* ⁵ who studied the prognosis of early-stage OA. They favored one outcome measure for functional status, so if a study reported multiple measures for functional status, performance measures were preferred above self-reported measures. And if performance measures were not applied, the results of the disease-specific, validated self-rated measures were reported. When more than one pain measure was used, the visual analogue scale (VAS) scores were reported.

The intervention trials of which the control group data was used, did not report statistical analyses on changes within the control group. Therefore, improvement or worsening of functional status and pain was determined to be present if 1) >60% of the patients reported such outcome on a dichotomous scale; 2) the difference between improved patients and patients whose functioning deteriorated, measured on an ordinal scale, was $\geq 10\%$; or 3) the difference between baseline and follow-up pain or functional status on a continuous scale was $\geq 10\%$. For the studies that did report proper statistical analyses, we used results presented by the authors. Moreover, to estimate the relative magnitude of change measured by different instruments, we calculated, when possible, effect sizes (ES) (difference between the inclusion and preoperative score divided by the standard deviation of inclusion scores). Effect sizes of 0.2, 0.5, and 0.8 are regarded as indicating small, medium, and large degrees of change, respectively ¹⁶. A negative ES represents deterioration, while a positive ES represents improvement.

The data were examined and found to be inappropriate for meta-analysis, as the included studies were considered heterogeneous with regard to study quality, analytic methods, design, covariates measured, outcome measures, and the length of waiting times. Therefore

we performed a qualitative data analysis (best-evidence synthesis) in which solely high quality (>80%) studies were reviewed ¹⁷. Levels of evidence were based on two earlier reviews of progress of hip and knee OA ^{5,18} (Table 1). Moreover, we differentiated between short (<180 days) and long (≥180 days) mean or median wait times.

Table 1. Levels of evidence that were applied in the best-evidence synthesis.

Level of evidence	
Strong evidence	Generally consistent findings in multiple high-quality cohort studies
Moderate evidence	Generally consistent findings in 1 high-quality cohort study and ≥2 high-quality case-control studies, or in ≥3 high-quality case-control studies
Limited evidence	(Generally consistent) findings in a single high-quality cohort study, or in ≤2 high-quality case-control studies
Conflicting evidence	Conflicting findings in high-quality studies (i.e., <75% of the studies reported consistent findings)
No evidence	No high-quality studies could be found

To examine the robustness of our results from our synthesis, two sensitivity analyses were performed: 1) by repeating the qualitative data analysis with 60% as a cut-off point for high methodological quality instead of 80%, and 2) by exclusively including adequately powered studies. Studies were considered adequately powered if the sample size was at least 110 patients; as 110 patients (allowing 10% attrition) would provide 80% power (assuming two-sided 5% significance) in patients waiting for THR to detect a clinically relevant difference (10%) in the WOMAC total score ¹⁹.

RESULTS

Search:

The bibliographic search retrieved 3426 articles, of which 17 articles - reporting on 15 studies - fulfilled all inclusion criteria ²⁰⁻³⁶ (Fig. 1). Because two eligible articles appeared identical, only published under a different title in a different journal one year apart ^{21,30}, we excluded the latter from the analysis ²¹. Moreover, McHugh *et al* published two eligible articles ^{20,27}, with different objectives, from the same set of data. We included the study for which objectives best matched our inclusion criteria ²⁷. Two studies, missed by the bibliographic search, were found through hand searching the references ^{37,38}, neither of these articles met the inclusion criteria. This left 15 eligible studies ²²⁻³⁶.

Study characteristics:

Of the 15 studies included, there were nine descriptive studies ²²⁻³⁰ and six RCTs ³¹⁻³⁶; for

RCTs with an intervention other than waiting, data were extracted from the control group, defined as the group receiving no intervention.

Ten studies assessed patients with end-stage hip OA^{23-29,32,34,35} and nine studies assessed patients with end-stage knee OA^{22,25-27,30-33,35}. The 15 articles together involved 1646 patients (mean age varied from 59 to 72 years with a range of 27-89), of which 788 and 858 patients waited for a new hip or knee respectively. The mean waiting period varied from 42 to 399 days (range 8-2486) (Table 2). None of the nine non-randomized studies assessed the patients' health status with performance measures.

Quality Appraisal:

The two reviewers scored 495 quality criteria and agreed on 408 (82%), leading to an interrater reliability of 0.64 assessed with Cohen's Kappa; disagreements were particularly found in the randomized trials. After one meeting, full consensus was reached.

Overall, quality was diverse (Table 3). Quality scores of the studies ranged from 10 - 28/33 on the Tooth checklist¹⁵. Two trials fulfilled more than 80% of the quality items^{25,26}, eight trials scored above 60%^{25-29,31-33}, and the remaining seven trials scored 60% or less^{22-24,30,34-36}. Four trials about hip OA were sufficiently powered^{24-26,29}; for knee OA two studies were able to refute the null-hypothesis^{26,31}.

End-stage hip osteoarthritis

Effects of a short term waiting period on functional status and pain

Two studies that described changes in pain and functional status during the waiting list period for THR were high quality^{25,26}. Kapstad *et al* (2007)²⁵ reported on the course of functional status and pain over a (median) period of 71 (range 21 to 362) days in 120 patients with a mean age (SD) of 67.4 (9.9) years. Kelly *et al* (2001)²⁶ reported over an average period of 132 days in 135 patients with a mean age (range) of 68.1 (27-89) years. Both studies described changes in the pain and function scales of the WOMAC-questionnaire while waiting. No statistically significant changes were found in pain or functional status, providing strong evidence that pain and self-rated functional status do not deteriorate in patients with end-stage OA waiting for THR. The studies both report that a proportion of the population did experience clinically relevant decline (defined as at least 10% worsening) in functional status (17%-35%) and pain (25%-27%) during the waiting list period. However, equivalent numbers are found of people that report improvement in pain (19%-26%) and functioning (14%-39%).

Effects of a long term waiting period on functional status and pain

No high quality studies examined the effect of mean long waiting times (≥ 180 days) on pain or functioning.

Table 2a. Overview of the characteristics of the included studies on hip osteoarthritis.

Study (period)	In- or exclusion criteria	Design	Data collection	Setting/ Country	Factors checked for association
Berge ³⁶ (??)	In: OA, on THR WL.	RCT [†]	Interview	Unclear (UK)	-
Chakravarty ²³ ('00-'03)	In: On THR WL.	Desc	(Postal) survey	1 surgeon 1 hospital (UK)	MHHS scores, age
Fielden ²⁴ ('99-'02)	In: ≥20y, OA, on THR WL. Ex: contra-lateral arthritis, hip dysplasia	Desc	(Postal) survey	4 hospitals (New Zealand)	-
Kapstad ²⁵ ('03-'04)	In: >18y, on THR WL.	Desc	Postal survey	6 hospitals in 3 counties (Norway)	DG, SE, WT, WOMAC
Kelly ²⁶ ('95-'97)	In: On THR WL. Ex: Revisions, surgery within one month.	Desc	Interview (Home visit)	24 surgeons at 2 Acute Care facilities(Canada)	DG, SE, BMI, WT, CM, SF36, WOMAC, meds
McGregor ³⁴ (??)	In: On primary THR WL. Ex: No bilateral THR, CM, or RA.	RCT [†]	Assessment	1 hospital (UK)	-
McHugh ²⁷ ('03, 3 mo)	In: >18y, OA, on THR WL.	Desc	Interview (at home)	1 orthopedic centre (UK)	-
Nilsdotter ²⁸ ('97-'98)	In: OA, on THR WL. Ex: Contra-lateral THR in 1 year, recurrent dislocations.	Desc	Assessment	1 hospital (Sweden)	-
Ostendorf ²⁹ ('99-'00)	In: >18y, on THR WL. Ex: RA, contra-lateral THR.	Desc	Postal survey	3 hospitals (The Netherlands)	DG, SE, WT, CM, previous surgery, walking ability
Rooks ³² ('01-'03)	In: OA, on THR WL. Ex: Inflammatory arthritis, Parkinson's, or severe CM.	RCT [†]	Assessment	7 orthopedic surgeon practices (US)	-

Abbreviations: [†] = control group data, BMI = Body Mass Index. CM = Comorbidity, DG = Demographic Info, Quality of Life, OHS =Oxford Hip Score, RA = Rheumatoid Arthritis, SE = Socio-economic information, SF36 = McMasters Universities Osteoarthritis Index, WT = Wait Time.

(continuation Table 2a)

Patients			Wait time	Conclusion of the authors (Main outcome measure)
N	♀%	Age		
21	79	71 (6.1)	120-180	NA (NPRS, 4-min walk)
94	-	-	225 (13-2486)	15% & 34% of the patients improve on pain and FS, others decline. (MHHS)
122	65	66 (35-85)	153 (30-510)	Longer WT incur greater economic costs and decline in PF. (WOMAC)
120	70	67.4 (9.9)	71 (31-362)	No change in pain or PF. (WOMAC)
135	59	68.1 (27-89)	131.5	No change in pain and dysfunction. (WOMAC, SF36)
20	-	72.8 (10.1)	408 (453)	NA (WOMAC, VAS)
48	59	68 (60-76)	180 (0)	Pain and PF deteriorated after 3 months on the WL. (WOMAC).
33	-	72 (52-86)	153 (120-300)	General health and pain improve, possibly due to regression to the mean. (WOMAC, SF36).
23	-	72 (52-86)	60 (30-90)	
143	65.8	68.4 (9.7)	180 (108)	Deterioration in pain and PF. (WOMAC, SF36, OHS).
31	52	59 (7)	70 (56-84)	NA (WOMAC, SF36, TUG)

FS = Functional Status, MHHS = Modified Harris Hip Score, PF = Physical Function, Qol = Short-Form-36, TUG = Timed up and go, WL = Waiting List, WOMAC = Western Ontario

Table 2b. Overview of the characteristics of the included studies on knee osteoarthritis.

Study	In- or exclusion criteria	Design	Data collection	Setting/ Country	Factors checked for association
Ahmad ²² (‘03-‘04)	In: None stated.	Desc	Interview	1 hospital (UK)	WT, baseline scores
Beaupre ³³ (??)	In: >40&<75y, OA, on TKR WL.	RCT [†]	Assessment	7 surgeons (Canada)	-
Hirvonen ³¹ (‘02-‘03)	In: >16y, on TKR WL. Ex: RA, fractures, hemophilia and deformity.	RCT	(Postal) survey	3 hospitals (Finland)	-
Kapstad ²⁵ (‘03-‘04)	In: >18y, on TKR WL.	Desc	Postal survey	6 hospitals in 3 counties (Norway)	DG, SE, WT, WOMAC
Kelly ²⁶ (‘95-‘97)	In: On TKR WL. Ex: Revisions, surgery within one month.	Desc	Interview (Home visit)	24 surgeons at 2 Acute Care facilities(Canada)	DG, SE, BMI, WT, CM, SF36, WOMAC, meds
McHugh ²⁷ (‘03, 3 mo)	In: >18y, OA, on TKR WL.	Desc	Interview (at home)	1 orthopedic centre (UK)	-
Pace ³⁰ (‘00-‘03)	In: >18y, OA. Ex: RA, other cause for knee pain.	Desc	Assessment	1 hospital (UK)	-
Rooks ³² (‘01-‘03)	In: OA, on TKR WL. Ex: Inflammatory arthritis, Parkinson’s, or severe CM.	RCT [†]	Assessment	7 orthopedic surgeon practices (US)	-
Weidenhielm ³⁵ (??)	In: OA, on TKR WL.	RCT [†]	Assessment	Unclear (Sweden)	-

Abbreviations: [†] = control group data, BMI = Body Mass Index. CM = Comorbidity, DG = Demographic Info, OKS = Oxford Knee Score, RA = Rheumatoid Arthritis, SE = Socio-economic information, SF36= Short-Form-36 Index, WT = Wait Time.

(continuation Table 2b).

Patients			Wait time	Conclusion of the authors (Main outcome measure)
N	♀% ♂%	Age		
58	51.7	68.7 (47-86)	242 (100-428)	FS in the majority of patients continues to deteriorate. (OKS)
66	50	67 (6)	42 (0)	NA (WOMAC, VAS)
127	70.1	66 (9.3)	73 (8-600)	Longer WT does not result in worse health related quality of life. (15D)
183	67.8	69 (9.0)	266 (28-818)	
50	80	69.2 (9.1)	102 (33-322)	No change in pain and decline in PF. (WOMAC)
178	59	68.1 (27-89)	131.5	No change in pain and dysfunction. (WOMAC, SF36)
57	59	68 (60-76)	90 (0)	Pain and PF deteriorated after 3 months on the WL. (WOMAC).
96	56	69 (49-80)	399 (273-660)	No deterioration in pain and PF. (OKS, AKSS)
23	52	59 (7)	70 (56-84)	NA (WOMAC, SF36, TUG)
20	-	-	90 (0)	NA (WOMAC, SF36, TUG)

FS = Functional Status, MHHS = Modified Harris Hip Score, PF = Physical Function, QoL = Quality of Life,
TUG = Timed up and go, WL = Waiting List, WOMAC = Western Ontario McMasters Universities Osteoarthritis

Table 3. Methodological quality score of each included study.

Studies (Author, Year of Publication)	MQ Items fulfilled (%)	Sufficiently powered (n≥110)	
		Hip	Knee
Kapstad <i>et al</i> , 2007 ²⁵	28 (85)	√	-
Kelly <i>et al</i> , 2001 ²⁶	27 (82)	√	√
Rooks <i>et al</i> , 2006 ³²	25 (76)	-	-
Hirvonen <i>et al</i> , 2007 ³¹	24 (73)	-	√
McHugh <i>et al</i> , 2007 ²⁷	24 (73)	-	-
Ostendorf <i>et al</i> , 2004 ²⁹	22 (67)	√	-
Nilsdotter <i>et al</i> , 2002 ²⁸	21 (64)	-	-
Beaupre <i>et al</i> , 2004 ³³	21 (64)	-	-
Fielden <i>et al</i> , 2005 ²⁴	18 (55)	√	-
Berge <i>et al</i> , 2004 ³⁶	17 (52)	-	-
Pace <i>et al</i> , 2005 ³⁰	14 (42)	-	-
Ahmad and Konduru, 2007 ²²	14 (42)	-	-
Weidenhielm <i>et al</i> , 1993 ³⁵	13 (39)	-	-
McGregor <i>et al</i> , 2004 ³⁴	11 (33)	-	-
Chakravarty <i>et al</i> , 2005 ²³	10 (30)	-	-

Abbreviations: MQ = Methodological quality

End-stage knee osteoarthritis

Effects of a short term waiting period on functional status and pain

Two studies that examined changes in pain and functional status during the waiting list period in patients with end-stage OA were high quality ^{25,26}. One study investigated the course of pain and functional status in 178 patients with a mean age (range) of 68.1 (27-89) years over an average period of 132 days ²⁶. The other study reported results on the course of pain and functional status in 50 patients with an average (range) age of 68.7 (47-86) years over a median waiting period of 102 (33-322) days ²⁵.

Ambiguous results are reported on self-rated functional status. Kapstad *et al* (2007) showed a small (ES=-0.20) but significant deterioration ²⁵, while Kelly *et al* (2001) reported no difference ²⁶ in self-rated functioning measured with the WOMAC, resulting in conflicting evidence. Results on pain, on the other hand, are similar, providing strong evidence that pain does not alter significantly while waiting for surgery. And, although, both authors ^{25,26} reported a clinically relevant decline (defined as at least 10% worsening) in a proportion of patients in pain (22%-25%) and self-rated functioning (20%-31%), a similar proportion of patients improved in pain (22%-30%) and functioning (10-37%) while waiting for surgery.

Effects of a long term waiting period on functional status and pain

No high quality studies examined the effect of long waiting times (≥180 days) on pain or functioning.

Prognostic factors

The aforementioned high quality trials also assessed the effect of prognostic variables on the change in pain and functional status. Kapstad *et al* (2007), using a multiple linear regression analysis, distinguished between hip (N=119) and knee (N=50) OA, while Kelly *et al* (2001) reported results of a logistic regression analysis in a combined sample of hip and knee OA (N=313). All investigated prognostic factors are shown in Table 4.

Table 4. Prognostic factors of change in self-reported functional status and pain in patients with hip or knee OA waiting less than six months.

	Protective factors:	Risk factors:	No association:	Unclear:
Change in self-reported function	<u>NE</u>	<u>NE</u>	<u>SE:</u> Wait time; age; marital status; female gender. <u>LE:</u> BMI; comorbidities; previous arthroplasty; education; working status; medication; social function; vitality; role emotional; mental health.	<u>CE:</u> WOMAC baseline scores pain and function.
Change in pain	<u>NE</u>	<u>SE:</u> Female gender	<u>SE:</u> Wait time; age; marital status. <u>LE:</u> BMI; comorbidities; previous arthroplasty; education; working status; medication; social function; vitality; role emotional; mental health.	<u>CE:</u> WOMAC baseline scores pain and function.

Abbreviations: BMI = Body Mass Index, CE = Conflicting Evidence, LE = Limited Evidence, NE = No evidence, SE = Strong evidence

Both studies found that women waiting for knee or hip replacement are at risk for intensified pain, therefore providing strong evidence. Moreover, strong evidence was found that wait time and age are not associated with the change in health status^{25,26}. The studies report conflicting evidence concerning the association of the baseline WOMAC pain and function scores on the change in these WOMAC-scales.

Sensitivity analyses

Methodological quality >60%:

Decreasing the 'high-quality' cut-off score from 80% to 60% yielded six additional studies^{27-29,31-33}. Five studies investigated patients with end-stage knee OA^{25-27,31,33} and five studies investigated patients with hip OA^{25-28,32}. For hip OA, this analysis complemented our results the effect of long waiting times, as it provided strong evidence that patients deteriorate on pain (Table 5). The sensitivity analyses provided conflicting results on the association between length of the waiting list and deterioration in pain and functioning, resulting in conflicting evidence.

Table 5. Results of the best evidence synthesis and the sensitivity analyses on studies describing functional status and pain in end-stage hip OA.

	Short waiting time (<180 days)		Long waiting time (≥180 days)	
	Pain	FS	Pain	FS
MQ>80%	No change (SE) ^{25,26}	No change (SE) ^{25,26}	No evidence (-)	No evidence (-)
MQ>60%	?? (CE) ^{25-28,32}	?? (CE) ^{25-28,32}	Deterioration (SE) ^{27,29}	?? (CE) ^{27,29}
N≥110	?? (CE) ²⁴⁻²⁶	No change (SE) ²⁴⁻²⁶	Deterioration (LE) ²⁹	Deterioration (LE) ²⁹

Abbreviations: FS = Functional Status, SE = Strong Evidence, CE = Conflicting Evidence, LE = Limited Evidence.

For knee OA, this analysis showed strong evidence supporting our finding that a short waiting period does not affect pain intensity, while it augmented our results on the effect of a long waiting list with limited evidence that pain might intensify (Table 6). Results on self-reported functional status all resulted in conflicting evidence. Our conclusions on the prognostic variables remained unaltered.

Table 6. Results of the best evidence synthesis and the sensitivity analysis on studies describing functional status and pain in end-stage knee OA.

	Short waiting time (<180 days)		Long waiting time (≥180 days)	
	Pain	FS	Pain	FS
MQ>80%	No change (SE) ^{25,26}	?? (CE) ^{25,26}	No evidence (-)	No evidence (-)
MQ>60%	No change (SE) ^{25-27,33}	?? (CE) ^{25-27,31,33}	Deterioration (LE) ²⁷	?? (CE) ^{27,31}
N≥110	No change (SE) ^{26,31}	No change (SE) ^{26,31}	No change (LE) ³¹	No change (LE) ³¹

Abbreviations: FS = Functional Status, SE = Strong Evidence, CE = Conflicting Evidence, LE = Limited Evidence.

Sufficiently powered studies:

When including sufficiently powered studies, five studies met the criteria^{24-26,29,31}. For hip OA^{24-26,29}, strong evidence showed no change in self-reported functional status, which strongly supports our previous findings. Moreover, we found limited evidence that self-reported functional status and pain deteriorate while waiting a long period (≥180 days) (Table 5). Ostendorf *et al* (2004) reported that the length of waiting time was a significant predictor for deterioration in scores, which is in disagreement with our earlier findings; resulting in conflicting evidence.

For knee OA ^{26,31}, no change was found in pain or self-reported physical function regardless of the duration of the waiting list (Table 6). Prognostic factors for future pain and functional status in knee OA remained unchanged, however the level of evidence decreased to moderate.

DISCUSSION

The results of this review indicate that whilst waiting a moderate time (<6 months) for joint replacement there is strong evidence that pain (hip and knee OA) and self-reported functional status (hip OA) do not change. There is conflicting evidence for change in self-rated functioning in patients with knee OA. Indefinite results are reported towards long waiting times (≥6 months), though pain might increase in people with hip OA. With regard to predictors of future pain or functional status, female gender was found to be the only predictor for intensified pain.

To our knowledge we are the first to systematically review and to utilize a best evidence synthesis on the literature describing the natural course of pain and functional status while waiting for elective major joint replacement. One other review, found in the grey literature, also reported on the matter ¹¹. They based their results on fewer included studies and did not perform a qualitative data analysis. They concluded that there is some evidence indicating that deterioration in functional health status occurs in patients waiting more than 6 months for joint replacement surgery. This conclusion is partially corroborated by the results of our review since the moderate quality studies, as included in the sensitivity analyses, reported mainly deterioration or conflicting evidence.

Implications of our review for current practice are twofold. First, according to the results of our review there is no need to prioritize patients for surgery, when 1. the waiting list is less than six months and 2. the reason for prioritization is to avoid further deterioration in pain or experienced functioning. However, one can think of legion other reasons for prioritizing patients. For example, to reduce the burden of living with severe levels of pain or to treat those patients unable to function independently due to the pain and disability associated with hip or knee osteoarthritis. Second, this review illustrates the lack of data on possible modifiable determinants of postoperative functional status (e.g. self-efficacy, (pain) coping strategies, fatigue, anxiety, physical activity and functional performance). Preoperative care, such as exercise and education could be of value to prepare patients before surgery to optimize the outcomes of surgery. However, reviews on preoperative care before joint replacement ³⁹⁻⁴¹ have not reported any clinically important effects. We believe that insight in the natural course of a disease is crucial in the development of proper, patient-specific, therapeutically relevant treatment programs. The studies included in this review mainly focused on the influence of demographic and socioeconomic data. We infer that this

is a hiatus in the current evidence. More insight into these determinants would allow clinicians to identify the patients at risk for functional decline, and consequently provide appropriate care to optimize the outcome of surgery and rehabilitation after surgery.

We believe there might be an under-estimation of the number of patients who experience deterioration whilst waiting for major joint replacement. Both high quality studies^{25,26} mentioned a subgroup experiencing deterioration; however an equally sized subgroup also experienced improvements over time. The notion is based on several indicators. First, the influence of ceiling effect in this respect has not yet been discussed. Patients with a very high baseline score are less likely to deteriorate on a questionnaire with a ceiling effect. Second, it is likely that the physicians in the different studies prioritized the patients of which they believed could deteriorate. This could also explain why the duration of the waiting list is not associated with future pain or functional status. Finally, regression towards the mean⁴² is often reported in the included studies to explain the improved pain and functioning in a subgroup of patients. However, none of the studies we included specifically tested their data for this kind of bias. To summarize, it is possible that the group of patients at risk for deterioration during the waiting list period is more significant than the results of this systematic review imply.

Drawing robust conclusions was hampered due to the pervasive use of self-reported measures in the included studies. We found, in agreement with Ethgen *et al* (2004)⁴³, that current research lacks the use of physical performance measures and seems to favor self-reported measures. Two major issues could arise when one's health status is investigated just with self-reported measures. First of all, a growing body of literature indicates that self-report measures of function provide different information than physical performance measures⁴⁴⁻⁴⁷; likely explained by the influence of pain on self-reported functioning⁴⁸. Therefore, we believe that clinically relevant deterioration before surgery in a patient's actual functional status or exercise capacity could contribute to the clinical decision-making of surgeons or physiotherapists. Secondly, results can easily be confounded by use of self-rated measures. Kapstad *et al* (2007) mentioned this phenomenon, known as response shift, as a possible explanation for their results²⁵. Response shift can bias findings because of changes over time regarding internal standards, values, or conceptualization of health related quality of life⁴⁹. This may result in an underestimation of the measured effects. In conclusion, the results of this review must be interpreted with care, as the results derived on two of the studies^{25,26} which only utilized self-rated measures.

A limitation of the review was our strict eligibility criteria. This led to the exclusion of two prospective trials^{50,51}. The study of Mahon *et al* (2002) was excluded since they only reported their waiting list data in histograms over the period from GP referral to surgery. The study of Hirvonen *et al* (2006) was excluded since they did not differentiate between hip and knee OA. To assess the implications of the exclusion of these studies on our results, we included both studies in our sensitivity analyses (data not shown). The inclusion of these studies did not affect the results of the sensitivity analyses.

The results of our review have some implications for future research. First, we believe, future studies should also include performance-based measures⁴⁴⁻⁴⁸ to properly investigate the progression of hip or knee OA during time spent on the waiting list. However, this recommendation automatically emphasizes the need for more high-quality studies that assess the measurement properties of performance-based methods in patients with hip or knee OA and the need for consensus on what activities should be included in a performance-based test, as well as the specific aspects of function should be measured⁴⁶. Second, these studies should investigate the role of confounders (eg. response shift bias) meticulously. And finally, studies should investigate more determinants of change in functional status (eg. self-efficacy), as current research mainly focused on the influence of demographic and socioeconomic data.

In conclusion, there is strong evidence that pain (hip and knee) and perceived functional status (hip) do not deteriorate in patients waiting for major joint replacement during a period less than six months. Conflicting evidence was found for the change in self-reported functional status while waiting less than six months for TKR. However, our conclusions have limited robustness given the small number of high quality studies performed in this field and the inconsistent results of the sensitivity analyses.

REFERENCE LIST

- (1) Crowninshield RD, Rosenberg AG, Sporer SM. Changing demographics of patients with total joint replacement. *Clin Orthop Relat Res.* 2006;443:266-272.
- (2) Ostendorf M, Johnell O, Malchau H, Dhert WJ, Schrijvers AJ, Verbout AJ. The epidemiology of total hip replacement in The Netherlands and Sweden: present status and future needs. *Acta Orthop Scand.* 2002;73:282-286.
- (3) Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am.* 2007;89:780-785.
- (4) Sharma L, Cahue S, Song J, Hayes K, Pai YC, Dunlop D. Physical functioning over three years in knee osteoarthritis: role of psychosocial, local mechanical, and neuromuscular factors. *Arthritis Rheum.* 2003;48:3359-3370.
- (5) van Dijk GM, Dekker J, Veenhof C, van den Ende CH. Course of functional status and pain in osteoarthritis of the hip or knee: a systematic review of the literature. *Arthritis Rheum.* 2006;55:779-785.
- (6) Garbuz DS, Xu M, Duncan CP, Masri BA, Sobolev B. Delays worsen quality of life outcome of primary total hip arthroplasty. *Clin Orthop Relat Res.* 2006;447:79-84.
- (7) Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am.* 2004;86-A:2179-2186.
- (8) Hajat S, Fitzpatrick R, Morris R, Reeves B, Rigge M, Williams O et al. Does waiting for total hip replacement matter? Prospective cohort study. *J Health Serv Res Policy.* 2002;7:19-25.
- (9) Lane NE. Clinical practice. Osteoarthritis of the hip. *N Engl J Med.* 2007;357:1413-1421.
- (10) Wylde V, Dieppe P, Hewlett S, Learmonth ID. Total knee replacement: Is it really an effective procedure for all? *Knee.* 2007;14:417-423.
- (11) Noseworthy, T. W., Sanmartin, C, Bohm, E, Connor-Spady, B, DeCoster, C, Dunbar, M et al. Toward Canadian Benchmarks for Health Service Wait Times - Evidence, Application and Research Priorities. 1-56. 14-10-2005.
Ref Type: Report
- (12) Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet.* 2001;357:1191-1194.
- (13) Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet.* 2002;359:57-61.
- (14) Sanmartin, C. Toward Standard Definitions of Waiting Times for Health Care Services. A Working Paper for the Western Canada Wait List Project. 2001.
Ref Type: Report
- (15) Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. *Am J Epidemiol.* 2005;161:280-288.
- (16) Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* 2 ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

- (17) Slavin RE. Best evidence synthesis: an intelligent alternative to meta-analysis. *J Clin Epidemiol.* 1995;48:9-18.
- (18) Lieveense AM, Bierma-Zeinstra SM, Verhagen AP, Verhaar JA, Koes BW. Prognostic factors of progress of hip osteoarthritis: a systematic review. *Arthritis Rheum.* 2002;47:556-562.
- (19) Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. *J Rheumatol.* 2000;27:2635-2641.
- (20) McHugh GA, Luker KA, Campbell M, Kay PR, Silman AJ. A longitudinal study exploring pain control, treatment and service provision for individuals with end-stage lower limb osteoarthritis. *Rheumatology (Oxford).* 2007;46:631-637.
- (21) Pace A, Orpen N, Doll H, Crawford EJ. Outcome scoring system evaluation of knee osteoarthritis in patients awaiting TKA. *J Knee Surg.* 2006;19:85-88.
- (22) Ahmad I, Konduru S. Change in functional status of patients whilst awaiting primary total knee arthroplasty. *Surgeon.* 2007;5:266-267.
- (23) Chakravarty D, Tang T, Vowler SL, Villar R. Waiting time for primary hip replacement--a matter of priority. *Ann R Coll Surg Engl.* 2005;87:269-273.
- (24) Fielden JM, Cumming JM, Horne JG, Devane PA, Slack A, Gallagher LM. Waiting for hip arthroplasty: economic costs and health outcomes. *J Arthroplasty.* 2005;20:990-997.
- (25) Kapstad H, Rustoen T, Hanestad BR, Moum T, Langeland N, Stavem K. Changes in pain, stiffness and physical function in patients with osteoarthritis waiting for hip or knee joint replacement surgery. *Osteoarthritis Cartilage.* 2007;15:837-843.
- (26) Kelly KD, Voaklander DC, Johnston DW, Newman SC, Suarez-Almazor ME. Change in pain and function while waiting for major joint arthroplasty. *J Arthroplasty.* 2001;16:351-359.
- (27) McHugh GA, Luker KA, Campbell M, Kay PR, Silman AJ. Pain, physical functioning and quality of life of individuals awaiting total joint replacement: a longitudinal study. *J Eval Clin Pract.* 2008;14:19-26.
- (28) Nilsson AK, Lohmander LS. Age and waiting time as predictors of outcome after total hip replacement for osteoarthritis. *Rheumatology (Oxford).* 2002;41:1261-1267.
- (29) Ostendorf M, Buskens E, van Stel H, Schrijvers A, Marting L, Dhert W et al. Waiting for total hip arthroplasty: avoidable loss in quality time and preventable deterioration. *J Arthroplasty.* 2004;19:302-309.
- (30) Pace A. The natural history of severe osteoarthritis of the knee in patients awaiting total knee arthroplasty. 2005.
- (31) Hirvonen J, Blom M, Tuominen U, Seitsalo S, Lehto M, Paavolainen P et al. Evaluating waiting time effect on health outcomes at admission: a prospective randomized study on patients with osteoarthritis of the knee joint. *J Eval Clin Pract.* 2007;13:728-733.
- (32) Rooks DS, Huang J, Bierbaum BE, Bolus SA, Rubano J, Connolly CE et al. Effect of preoperative exercise on measures of functional status in men and women undergoing total hip and knee arthroplasty. *Arthritis Rheum.* 2006;55:700-708.

- (33) Beaupre LA, Lier D, Davies DM, Johnston DB. The effect of a preoperative exercise and education program on functional recovery, health related quality of life, and health service utilization following primary total knee arthroplasty. *J Rheumatol*. 2004;31:1166-1173.
- (34) McGregor AH, Rylands H, Owen A, Dore CJ, Hughes SP. Does preoperative hip rehabilitation advice improve recovery and patient satisfaction? *J Arthroplasty*. 2004;19:464-468.
- (35) Weidenhielm L, Mattsson E, Brostrom LA, Wersall-Robertsson E. Effect of preoperative physiotherapy in unicompartmental prosthetic knee replacement. *Scand J Rehabil Med*. 1993;25:33-39.
- (36) Berge DJ, Dolin SJ, Williams AC, Harman R. Pre-operative and post-operative effect of a pain management programme prior to total hip replacement: a randomized controlled trial. *Pain*. 2004;110:33-39.
- (37) Kili S, Wright I, Jones RS. Change in Harris hip score in patients on the waiting list for total hip replacement. *Ann R Coll Surg Engl*. 2003;85:269-271.
- (38) Derrett S, Paul C, Morris JM. Waiting for elective surgery: effects on health-related quality of life. *Int J Qual Health Care*. 1999;11:47-57.
- (39) Ackerman IN, Bennell KL. Does pre-operative physiotherapy improve outcomes from lower limb joint replacement surgery? A systematic review. *Aust J Physiother*. 2004;50:25-30.
- (40) Coudeyre E, Jardin C, Givron P, Ribinik P, Revel M, Rannou F. Could preoperative rehabilitation modify postoperative outcomes after total hip and knee arthroplasty? Elaboration of French clinical practice guidelines. *Ann Readapt Med Phys*. 2007;50:189-197.
- (41) McDonald S, Hetrick S, Green S. Pre-operative education for hip or knee replacement. *Cochrane Database Syst Rev*. 2004;CD003526.
- (42) Bland JM, Altman DG. Regression towards the mean. *BMJ*. 1994;308:1499.
- (43) Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. *J Bone Joint Surg Am*. 2004;86-A:963-974.
- (44) Maly MR, Costigan PA, Olney SJ. Determinants of self-report outcome measures in people with knee osteoarthritis. *Arch Phys Med Rehabil*. 2006;87:96-104.
- (45) Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskelet Disord*. 2005;6:3.
- (46) Terwee CB, Mokkink LB, Steultjens MP, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology (Oxford)*. 2006;45:890-902.
- (47) Kennedy DM, Stratford PW, Hanna SE, Wessel J, Gollish JD. Modeling early recovery of physical function following hip and knee arthroplasty. *BMC Musculoskelet Disord*. 2006;7:100.
- (48) van den Akker-Scheek I, Zijlstra W, Groothoff JW, Bulstra SK, Stevens M. Physical functioning before and after total hip arthroplasty: perception and performance. *Phys Ther*. 2008;88:712-719.

- (49) Schwartz CE, Andresen EM, Nosek MA, Krahn GL. Response shift theory: important implications for measuring quality of life in people with disability. *Arch Phys Med Rehabil.* 2007;88:529-536.
- (50) Mahon JL, Bourne RB, Rorabeck CH, Feeny DH, Stitt L, Webster-Bogaert S. Health-related quality of life and mobility of patients awaiting elective total hip arthroplasty: a prospective study. *CMAJ.* 2002;167:1115-1121.
- (51) Hirvonen J, Blom M, Tuominen U, Seitsalo S, Lehto M, Paavolainen P et al. Health-related quality of life in patients waiting for major joint replacement. A comparison between patients and population controls. *Health Qual Life Outcomes.* 2006;4:3.

6

Prevalence and predictors of health care use
in patients with early hip or knee
osteoarthritis: two-year follow-up data from
the CHECK cohort.

ABSTRACT

Objective. To describe health care utilization (HCU) and predict analgesic use and health professional (HP) contact at baseline and two years in individuals with early symptomatic hip and/or knee osteoarthritis (OA).

Design. Baseline and two-year data on HCU of the 1002 participants from the multi-centre Cohort Hip & Cohort Knee study were used. Six forms of health care services were described: analgesic use, supplement use, contact with a General Practitioner (GP), contact with a Health Professional (HP), contact in secondary care, and alternative medicine use. Multivariable logistic regression was performed in order to identify predisposing, enabling and disease-related variables that predict analgesic use and HP-contact at two years; treatment modalities of first choice in early OA.

Results. For the hip (n=170), the knee (n=414) and the hip and knee (n=418) group analgesic use (38%, 29% and 47%, respectively), contact with a GP (32%, 38% and 36%, respectively) and contact with a HP (26%, 18% and 20%, respectively), were reported most often at baseline. Contact with a GP significantly decreased, supplement use increased (to about one third), and other treatment modalities remained stable at two years. In all three groups, analgesic use at baseline was the strongest predictor for analgesic use at two years, whereas contact with a HP at baseline was the strongest predictor of contact with a HP after two years. Belonging to a first generation minority was a predisposing risk factor (Odds Ratio [95%-CI], 8.72 [1.55-48.97]) for analgesic use in the hip and knee group.

Conclusions. In early OA, familiarity with HCU and other predisposing factors are, apart from disease-related factors strongly associated with HCU at two years. Further research is necessary to examine whether our findings reflect sub-optimal management of early OA in terms of efficacy and equity.

Hoogeboom TJ, Snijders GF, Cats HA, de Bie RA, Bierma-Zeinstra SMA, van den Hoogen FHJ, van Riel PLCM, Emans PJ, Wesseling J, Broeder AA, van den Ende CHM.

Prevalence and predictors of health care use in patients with early hip or knee osteoarthritis: two-year follow-up data from the CHECK cohort.

Osteoarthritis and Cartilage 2012;20(6):525-31.

INTRODUCTION

Individuals with osteoarthritis (OA) often require long-term access to a broad range of health care services [1;2]. Numerous studies have shown high variability in the amount of health care utilization (HCU) in people with OA [3-13]. Moreover, a number of characteristics have been found to be cross-sectionally associated with higher HCU in people with OA, including obesity, being single, higher education level, pain, disabilities, depression, quality of life, comorbidities and previous HCU [3;4;7-12;14].

In current practice, OA is often under-diagnosed and under-treated [14;15] and patients who seek care early in their symptoms might simply be told there is nothing to be done, or that their symptoms are just part of the aging process. In our opinion, however, patients in the earliest phase of the disease could and should be adequately guided throughout the plethora of different treatment options [16]. Therefore, insight into factors that predict use of the health care system over time is important to allow health care providers to intervene and possibly optimize patients' use of health care resources. To date, however, longitudinal research on OA-related HCU is scarce [17-19] and no data on predictors of future HCU in individuals with OA are available. Another hiatus in the current body of HCU evidence is the lack of studies describing health care consumption of individuals with early OA.

Use of analgesics and contact with a health professional is recommended in the early phase of OA [16] and should be prescribed in consonance with the patients' needs. Andersen and Newman state that access to health care depends on three types of factors: predisposing, enabling, and disease-related factors [20;21]. Predisposing factors refer to demographic and social characteristics. Enabling resources reflect the ability to use care resources. Disease-related factors represent the most immediate cause for HCU, reflected by diagnosis, perception of illness, presence of symptoms, and disability [22]. We hypothesize that both the use of analgesics and contact with a health professional (HP) are the most utilized treatment options and can be predicted over time based on of disease-related factors in patients with early symptomatic OA.

Therefore, the aims of this study were: 1) to describe HCU, and 2) to identify predictors for future use of analgesics and contact with a HP in individuals with early symptomatic hip and/or knee OA. To do so, we used baseline and two-year follow-up data from the Cohort Hip & Cohort Knee (CHECK) study [23].

METHODS

Design

Baseline and two-year follow-up data were used from the CHECK cohort [23]. CHECK is a

prospective cohort study of 1002 individuals with early symptomatic OA of hip or knee in The Netherlands. These individuals will be followed prospectively for a total period of 10 years. A total of 10 general and university hospitals are participating in the study. Eligibility was determined by physicians of participating centres. Study visits for all participants were planned at baseline and at two, five, eight and 10 years and consist of structured interview, physical examination, radiological assessment, serum- and urine-analysis, and questionnaires.

The study was approved by the medical ethics committees of all participating centres, and all participants gave their written informed consent. Data on self-report and physical examination at baseline and two-year data on HCU from the CHECK cohort were used for the current study.

Study population

On entry, all participants had pain or stiffness of hip or knee, and were aged 45-65 years. All patients who visited the GP on their own initiative and who potentially met the inclusion criteria were referred to one of the 10 participating centres. Additionally, participants were recruited through advertisements and articles in the local newspapers and on the Dutch Arthritis Association (DAA)-website. Participants with any other pathological condition that could explain the symptoms were excluded (e.g. other rheumatic disease, previous hip or knee joint replacement, congenital dysplasia, osteochondritis dissecans, intra-articular fractures, septic arthritis, Perthes' disease, ligament or meniscus damage, plica syndrome, Baker's cyst) [23].

At baseline, the majority of this cohort (83%) reported knee symptoms [23], of whom 76% fulfilled the clinical ACR criteria for the classification of knee OA [24]. Hip symptoms were reported by 59% of the cohort [23], of whom 24% fulfilled the clinical ACR criteria for the classification of hip OA [25]. Two-year follow-up data show an increase in these percentages, as well as an increase in radiological signs. Therefore the CHECK cohort can be considered as an "early" symptomatic OA cohort [26].

Measurement instruments

Health care utilization (dependent variable):

An HCU questionnaire was developed based on the one developed by Patient Panel Chronic Diseases (The Netherlands Institute for Health Services Research; NIVEL) [27], and the questionnaire Economic Aspects in Rheumatoid Arthritis [28]. Participants reported whether or not they had visited health care providers during the last three months or were hospitalized during the last year for their hip and/or knee problems. All available OA-related health care services were included: the General Practitioner (GP), medical specialists (e.g. rheumatologist, orthopaedic surgeon), HP (i.e. physiotherapy, occupational therapy, exercise therapy, psychology), hospital stay, use of analgesics (i.e. prescription and

over-the-counter drugs) and/or supplements (i.e. glucosamine and/or chondroitin), and complementary and alternative medicine (CAM) (e.g. acupuncture, shiatsu massage, contact with a chiropractor, homeopathic medicine, etc).

Baseline predictors (independent variables):

Patient characteristic (sociodemographic data, lifestyle factors) and comorbidity were collected by use of a standardized questionnaire.

Patient-reported outcome measures were determined by use of validated questionnaires. Pain over the last week was measured with a numeric rating scale (NRS, 0-10) [29]. Condition specific health status was evaluated with the WOMAC [30;31]. WOMAC evaluates three dimensions, Pain (0 –20), Stiffness (0–8), and Functioning (0–68), where higher scores represent worse health status. Participants completed one WOMAC questionnaire for the hips and knees overall. Self-reported health-related quality of life was measured using the Short-Form 36 (SF-36) health survey [32]. The SF-36 consists of eight subscales with a score range of 0–100, where a score of 100 represents the best possible health situation. The physical (PCS) component summary scores was calculated and fatigue and distress were assessed with the vitality and mental health subscale of the SF-36, respectively [32]. Coping behaviour was measured with the Pain Coping Inventory [33]. Active coping was defined as the mean of three active strategies (pain transformation, distraction, and reducing demands) and passive coping as the mean of three passive strategy scores (retreating, worrying and resting). Value of own health was assessed with the visual analogue scale for health of the EuroQoL-5D questionnaire [34]. Social support was measured with the Dutch Social Support Scale [35].

Physical examination of the hip consisted of measuring range of motion (internal rotation and flexion, in degrees), and pain (yes/no) during internal rotation and flexion. For the knee physical examination consisted of range of motion (flexion-extension), pain during flexion, bony tenderness, effusion of the knee (refill test), crepitus, palpable warmth, and bony enlargement. The highest Kellgren-Lawrence Grading Scale score (range 0 to 4) of the two hips and two knees [36], was used as indicator for radiological severity.

In Box 1 all baseline predictors are depicted according to the classification of Andersen and Newman [20;21], namely predisposing factors, enabling factors, and disease-related factors.

Statistical analysis

We performed all statistical analyses using STATA/IC 12.0. A hip only group (n = 170), knee only group (n = 414) and hip and knee group (n = 418) were defined, based on self-reported joint pain by the participants during the first study visit. We described HCU at baseline and at two years separately for each group. Differences between baseline and two years were tested by use of McNemar tests and differences between groups by use of Kruskal-Wallis test. Complete-case data were used for all descriptive analyses.

Box 1. Baseline predictors categorized according to the model of Andersen and Newman.

Predisposing factors

Age; Sex; BMI; Ethnicity; Marital State; Education; Family Size, Work; Smoking; Alcohol consumption; Coping Style (Active & Passive); Previous HCU.

Enabling factors

Dependency on others; Social Support; Health insurance

Disease-related factors

Number of comorbidities; Pain during last week; Number of painful joints (hip/knee); WOMAC pain; WOMAC stiffness, WOMAC functioning; EQ-vas; SF-36 mental health; SF-36 vitality; SF-36 Physical Component Score; ESR.

Hip specific: Knee pain; ROM (Internal rotation); ROM (Flexion); Pain during internal rotation; Pain during flexion; Highest K&L grade hip.

Knee specific: Hip pain; Palpable warmth; Bony tenderness; Effusion; Bony enlargement; Crepitus; ROM (Flexion + Extension); Pain during flexion; Highest K&L grade knee.

Abbreviations: BMI, Body Mass Index; EQ-vas, EuroQol Visual Analogue Scale; ESR, Erythrocyte sedimentation rate; HCU, Health Care Utilization; K&L, Kellgren and Lawrence; ROM, Range of motion; SF-36, Short-Form 36, WOMAC, Western Ontario and McMaster University Osteoarthritis index.

To predict use of analgesics and contact with a HP in the hip only, the knee only and hip and knee group at two years, we used multivariable logistic modelling. Use of analgesics (yes or no) comprised both prescription and over-the-counter drugs. Contact with a HP (yes or no) was defined as being in contact with a physiotherapist, occupational therapist, or exercise therapist. First, we studied the individual bivariate association of each independent variable with the dependent variable by calculation of odds ratios. Variables with a p -value ≥ 0.157 [37] were removed and the remaining variables were tested for collinearity by use of the Variance Inflation Factor (VIF, cut-off >10) statistic [37]. Subsequently, multivariable logistic regression analyses were performed for the remaining predisposing, enabling and disease-related factors (p -value for removal < 0.20). Finally, we entered the remaining predictor variables as one block into a backward stepwise regression model (p removal 0.10) to come to the final model. Finally, Hosmer and Lemeshow goodness of fit test, Area Under the Receiver Operating Curve (AUC), and Nagelkerke's R^2 statistic were calculated.

Logistic regression models with indicator variables for missing values as outcome and the other variables as covariates [38] showed that missing data were not associated to observed values, indicating that imputation of the missing values may reduce bias and increase efficiency [39;40]. Multiple imputation using Imputation by Chained Equation was used to estimate missing values [41]. A total of 10 different multiply imputed datasets were generated and finally combined using Rubin's rule of combination [42]. All logistic regression analyses were performed on multiply imputed datasets.

RESULTS

Study sample

Of the 1002 participants included in the CHECK cohort, 170 participants reported only hip complaints, 414 only knee complaints and 418 both hip and knee complaints. A total of 982 (98%) participants completed and returned the baseline questionnaire and 932 (93%) participants returned the health service questionnaire after two years. Seventy-four percent of the participants had complete data and 90 percent had less than three missing values. Knee range of motion and erythrocyte sedimentation rate (ESR) were the independent variables with the largest number of missing values; 9% and 5%, respectively. Characteristics of the study sample are presented in Table 1; for a more detailed description of the total population, see Wesseling *et al* (2008) [23].

Table 1. Characteristics at baseline, presented for the hip only, knee only and, hip and knee group.

	Hip only group	Knee only group	Hip and knee group
<i>N</i>	170	414	418
Age (mean (SD))	55.7 (5.6)	56.1 (5.1)	55.8 (5.2)
Female	129 (76%)	317 (77%)	346 (83%)
BMI (median (IQR))	25.2 (22.7 – 27.8)	25.6 (23.6 – 28.5)	25.5 (23.5 – 28.4)
Married/Partnership	137 (82%)	334 (83%)	340 (83%)
Education			
Primary	2 (1%)	14 (3%)	9 (2%)
Secondary	111 (67%)	275 (69%)	296 (73%)
Higher Education	53 (32%)	112 (28%)	102 (25%)
Highest K&L score			
0	106 (63%)	247 (60%)	239 (57%)
1	37 (22%)	119 (29%)	139 (33%)
2	21 (13%)	41 (10%)	33 (8%)
3	4 (2%)	4 (1%)	6 (1%)
4	-	-	-
WOMAC (median (IQR))			
Pain	4.8 (3.2)	4.5 (3.4)	5.7 (3.5)
Stiffness	2.5 (1.7)	2.5 (1.7)	2.9 (1.7)
Functioning	14.7 (11.1)	14.2 (11.0)	18.3 (12.2)

Abbreviations: -, Exclusion criteria; BMI, body mass index; IQR, Inter-quartile Range; K&L, Kellgren-Lawrence Grading Scale; SD, Standard Deviation; WOMAC, Western Ontario and McMaster Universities osteoarthritis index.

Health care utilization (HCU)

The most frequently reported forms of health care use in the hip only, the knee only and the hip and knee group at baseline were analgesic use (38%, 29% and 47%, respectively),

contact with a GP (32%, 38% and 36%, respectively) and contact with a HP (26%, 18% and 20%, respectively). In all groups, contact with a GP significantly decreased and supplement use significantly increased at two years (all $p<0.05$). For other health care modalities (i.e. secondary care, contact with a HP, analgesics and CAM) a standstill in use was observed at two years in the hip only group and hip and knee group compared with baseline. In the knee only group HP contact and secondary care use were significantly decreased at two years (Table 2).

Compared with the other groups, participants in the hip and knee group reported significantly more HCU (in particular use of analgesics) at baseline and at two years. Between the hip only group and knee only group no statistically significant differences were found (Table 2).

Table 2. Health care utilization at baseline and two years for the hip only, the knee only, and the hip and knee group.^{\$}

	Hip only group (n=170)			Knee only group (n=414)			Hip and knee group (n=418)		
	HC U t0	HCU t2	No HCU t0&2	HCU t0	HCU t2	No HCU t0&2	HCU t0	HCU t2	No HCU t0&2
Contact GP	32%	10%	61%	38%	11%	58%	36%	17%	57%
Contact secondary care	11%	11%	79%	14%	5%	84%	12%	8%	81%
Contact HP	26%	17%	66%	18%	12%	75%	20%	21% [†]	66%
Analgesics	38%	39%	52%	29%	31%	58%	47% [‡]	50% [‡]	38%
Supplements*	18%	33%	62%	13%	31%	56%	17%	33%	62%
CAM	13%	8%	85%	6%	6%	91%	14% [†]	12%	80%

^{\$} Numbers printed bold are significantly different. * Glucosamine / Chondroitin. [†] Significantly greater than the knee group. [‡] Significantly greater than both the hip and knee group.

Abbreviations: CAM = Complementary and Alternative Medicine; GP = General Practitioner; HCU Health Care Utilization; HP = Health Provider; T0 = Baseline, T2 = at 2 years.

Predictors for analgesic use

For the hip only group, multivariable logistic modelling yielded two factors (one predisposing and one disease-related factor) that independently ($p<0.05$) predicted analgesic use at two years (Nagelkerke's R^2 : 0.47) (Table 3). Having used analgesics before and experiencing more problems in functioning (WOMAC) were both significant risk factors for analgesic use at two years.

For the knee only group, multivariable logistic modelling resulted in one predisposing, one enabling, and one disease-related factor ($p < 0.05$) that independently predicted analgesic use at two years (Nagelkerke's R^2 : 0.35) (Table 3).

Table 3. Predictors for analgesics use for OA-complaints at two years.

<i>N</i>	Hip only group	Knee only group	Hip+Knee group
	170	414	418
Predisposing factors			
Paid work, yes	-	0.64 [0.38,1.09]	-
Active coping	-	-	2.45** [1.36,4.42]
Ethnicity			
Majority	-	-	(reference)
1 st generation immigrant	-	-	8.72* [1.55,48.97]
2 nd generation immigrant	-	-	0.87 [0.28,2.67]
Analgesic use, t0	8.43*** [3.67,19.39]	4.69*** [2.73,8.05]	5.33*** [3.16,9.01]
Enabling factors			
Help by others, yes	-	7.08** [1.81,27.78]	-
Health Insurance			
Public	-	(reference)	-
Private	-	0.87 [0.33,2.30]	-
Private (government)	-	0.55 [0.31,1.00]	-
Disease-related factors			
No. of joint complaints (0-4)	NA	NA	1.30* [1.02,1.67]
WOMAC Pain	-	1.18*** [1.08,1.28]	-
WOMAC Functioning	1.07** [1.03,1.12]	-	-
SF-36 PCS	-	-	0.97** [0.94,0.99]
Bony Tenderness	NA	-	1.76* [1.08,2.85]
Hip pain during int. rotation	2.70 [0.93,7.83]	NA	-
Hip flexion	-	NA	2.16* [1.18,3.95]
Post-estimation results			
<i>AUC</i>	0.85 [0.78,0.92]	0.80 [0.75, 0.85]	0.83 [0.79,0.87]
<i>Nagelkerke's R²</i>	0.472	0.345	0.403
<i>Hosmer – Lemeshow statistic</i>	6.84, $p > 0.05$	3.93, $p > 0.05$	7.57, $p > 0.05$

Odds Ratio; 95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Abbreviations: AUC = Area Under the receiver operating characteristic Curve; NA = Not Applicable, No. = Number of; OA = Osteoarthritis; PCS = Physical Component Scale; SF-36 = Short-Form; t0 = baseline; WOMAC = Western Ontario McMasters University Osteoarthritis Index.

Risk factors for analgesic use were previous analgesic use, dependency on others and having a higher WOMAC pain score.

For the hip and knee group we identified two predisposing, one enabling and four disease-related factors ($p < 0.05$) that independently predicted analgesic use at two years (Nagelkerke's R^2 : 0.40). Risk factors were, previous analgesic use, having an active coping

style, belonging to a first generation immigrant, having a higher number of painful joints, experiencing bony tenderness and having reduced hip flexion (Table 3).

Table 4. Predictors for contact with a health professional at two years.

	Hip only group	Knee only group	Hip and knee group
<i>N</i>	170	414	418
Predisposing factors			
Contact with HP, t0	3.47 ^{**} [1.43,8.42]	4.15 ^{***} [2.04,8.47]	2.82 ^{***} [1.61,4.95]
Education, high	-	-	0.52 [*] [0.27,0.99]
Enabling factors			
-	-	-	-
Disease-related factors			
No. of comorbidities (0-5)	-	-	1.20 [0.99,1.46]
Pain during last week (VAS)	-	0.84 [0.69,1.02]	-
WOMAC Stiffness	-	1.22 [0.97,1.53]	-
SF-36 PCS	-	0.96 [0.93,1.01]	-
SF-36 Mental Health	-	0.97 ^{**} [0.95,0.99]	-
Knee refill	-	2.67 [*] [1.09,6.50]	-
Post-estimation results			
<i>AUC</i>	0.64 [0.53,0.74]	0.75 [0.68,0.83]	0.67 [0.60,0.73]
<i>Nagelkerke's R²</i>	0.079	0.169	0.090
<i>Hosmer – Lemeshow statistic</i>	N/A	5.92, p>0.05	3.22, p>0.05

Odds Ratio; 95% confidence intervals in brackets. ^{*} $p < 0.05$, ^{**} $p < 0.01$, ^{***} $p < 0.001$.

Abbreviations; AUC = Area Under the receiver operating characteristic Curve; HP = Health Professional; No. = Number of; OA = Osteoarthritis; SF-36 = Short-Form; t0 = baseline; WOMAC = Western Ontario McMasters University Osteoarthritis Index.

Predictors for health professional use

For the hip only group, one predisposing factor - previous contact with a HP - ($p < 0.05$) was identified to independently predict health professional care use at two years (Nagelkerke's R^2 : 0.08) (Table 4).

For the knee only group, one predisposing and two disease-related factors ($p < 0.05$) were identified for contact with a HP at two years (Nagelkerke's R^2 : 0.17) (Table 4). Risk factors for contact with a HP were: previous contact and a positive knee refill test. Lower level of distress (SF-36 mental health) was a protective factor.

For the hip and knee group, two predisposing factors were identified for contact with a HP at two years (Nagelkerke's R^2 : 0.09). Previous contact with a HP was a risk factor for contact with a HP at two years. Having a higher degree of education was a protective factor (Table 4).

DISCUSSION

In this longitudinal study we found an overall low use of HCU for the participants with early symptomatic hip and/or knee OA. At two years, supplement and analgesic use were most often reported. We were unable to confirm our hypothesis that HCU at two years was primarily explained by disease-related factors, as the strongest predictor for analgesic use and contact with a HP at two years was previous use of these care modalities. In addition, several other predisposing and enabling factors were identified which were associated with analgesic use or HP contact at two years.

As far as we know, we are the first to identify baseline predictors for future analgesic use and contact with a HP in individuals with early hip or knee OA. Others have described health care consumption over time in patients with (symptomatic) hip or knee OA [17-19]. These studies generally reported substantially more health consumption in people with OA than found in our study. These differences are likely explained by differences in time frame of the studies, sampling methods and management of OA among different countries.

An interesting finding in our study was that supplement use (i.e. glucosamine and chondroitin) was the only treatment modality that had increased after two-year period. The CHECK cohort was formed from 2002 to 2005 and glucosamine was registered in The Netherlands as an over-the-counter drug in 2005, although it was available before that. This registration, however, hardly increased overall supplement use during that period in The Netherlands. Another hypothesis for this increase could be that patients try other, more alternative forms of care after experiencing disappointing results from conventional treatments. However, this explanation is questionable, since other forms of alternative care had decreased at two years. A final hypothesis could be that GPs dismiss their patients' OA symptoms as mere pains and aches of aging and therefore would not embark on a well-thought out, individualized treatment plan. The latter was also (partly) demonstrated by the decrease/standstill in HCU after two years. Lack of a clear treatment plan could leave patients unattended in their needs and as a result leave patients searching for easily accessible forms of health care (e.g. supplements). Additional long-term observation of supplement use in the CHECK cohort will reveal whether this increase is persistent.

We identified that being a first generation minority was a risk factor for analgesic use in the hip and knee group, but not in the hip only and knee only group. Literature findings demonstrate that minority groups with osteoarthritis more frequently use over-the-counter drugs than Caucasian people [43]. At this moment we have no satisfactory explanation for the differences in our study between subgroups of patients. However, in our view our findings suggest potential racial disparities in medication use that warrant further research.

Our findings give reason to re-evaluate the current practice in people with OA, as we believe there is room for improvement. First, supplement use was the only health care modality that had increased at two years and approximately 10% of the participants

reported the use of CAM in earliest stage of the disease, whilst the effectiveness of such interventions are under debate [44]. Using these forms of therapy will most likely result in high therapy costs, disappointing therapy results and (unnecessary) risks for adverse events. Second, as mentioned in the previous paragraph, people belonging to minority groups might receive sub-optimal prescription of analgesics, indicating inequity of health care prescription in the management of OA. Third and final, our results suggest that factors other than disease-related factors (e.g. previous use of a therapy modality) are associated with contacts with HPs or use of analgesics, indicating that health status is not the primary determinant associated with these treatment modalities. Considering these three issues, we believe that patients should be guided throughout the plethora of health care modalities. Some potentially useful projects aiming to deliver optimal primary care have already been initiated, such as the Beating Osteoarthritis [16] and the MOSAICS project (ISRCTN06984617).

The strengths of our study include the large sample size, the participation of 10 different centres and the high response and retention rates. In addition, the amount of missing data was small. Limitations include the reliance on self-report data with regard to HCU which may have induced recall bias and most likely resulted in over-reporting [45]. Also, participants may have perceived the contact with the study assessor (physiotherapist or clinician) as health care consumption, which could also have resulted in a over-reporting of health care contacts. A second limitation is that participants in the CHECK cohort may have regarded the CHECK visits as a substitute for contacts with other care providers, thus possibly resulting in less utilization of health care. Finally, the cohort is limited to The Netherlands, which affects the study's generalizability due to its specific health care system. In The Netherlands, GPs are accessible to everyone since basic health insurance is mandatory. Dutch primary care, with gate-keeping GPs at its core, prevents unnecessary use of more expensive secondary care and promotes consistency and coordination of individual care. It is likely that the results in this study are particularly generalizable to countries with a similar health care system, such as the United Kingdom, Spain, Finland and Italy.

In conclusion, individuals with early symptomatic OA rely in the first two years mainly on analgesics, with approximately one third trying supplements at two years. We demonstrated that predisposing factors (such as education, ethnicity and familiarity with care) are strongly associated with HCU, whereas one would presume disease-related factors to be leading in the utilization of care. The latter suggests that HCU should be optimized by taking the health needs of patients in OA into consideration and by minimizing the influence of predisposing factors.

ACKNOWLEDGEMENTS

The authors thank the steering committee of the Dutch Arthritis Association, which is chaired by Professor J. W. J. Bijlsma and coordinated by J. Wesseling, MSc. Involved in the cohort are the Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen/Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede/Ziekenhuisgroep Twente Almelo; Reade, formerly Jan van Breemen Institute/VU Medical Center Amsterdam; Sint Maartenskliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen.

REFERENCE LIST

- (1) Hagglund KJ, Clark MJ, Hilton SA, Hewett JE. Access to healthcare services among persons with osteoarthritis and rheumatoid arthritis. *Am J Phys Med Rehabil* 2005;84:702-11.
- (2) Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. *Ann Rheum Dis* 2001;60:91-7.
- (3) Cisternas MG, Yelin E, Katz JN, Solomon DH, Wright EA, Losina E. Ambulatory visit utilization in a national, population-based sample of adults with osteoarthritis. *Arthritis Rheum* 2009;61:1694-703.
- (4) Cronan TA, Shaw WS, Gallagher RA, Weisman M. Predicting health care use among older osteoarthritis patients in an HMO. *Arthritis Care Res* 1995;8:66-72.
- (5) Dominick KL, Ahern FM, Gold CH, Heller DA. Health-related quality of life and health service use among older adults with osteoarthritis. *Arthritis Rheum* 2004;51:326-31.
- (6) Grindrod KA, Marra CA, Colley L, Cibere J, Tsuyuki RT, Esdaile JM, et al. After patients are diagnosed with knee osteoarthritis, what do they do? *Arthritis Care Res (Hoboken)* 2010;62:510-5.
- (7) Hopman-Rock M, de Bock GH, Bijlsma JW, Springer MP, Hofman A, Kraaijmaat FW. The pattern of health care utilization of elderly people with arthritic pain in the hip or knee. *Int J Qual Health Care* 1997;9:129-37.
- (8) Mitchell HL, Carr AJ, Scott DL. The management of knee pain in primary care: factors associated with consulting the GP and referrals to secondary care. *Rheumatology (Oxford)* 2006;45:771-6.
- (9) Rosemann T, Joos S, Szecsenyi J, Laux G, Wensing M. Health service utilization patterns of primary care patients with osteoarthritis. *BMC Health Serv Res* 2007;7:169.
- (10) Rosemann T, Laux G, Szecsenyi J. Osteoarthritis: quality of life, comorbidities, medication and health service utilization assessed in a large sample of primary care patients. *J Orthop Surg Res* 2007;2:12.
- (11) Rosemann T, Grol R, Herman K, Wensing M, Szecsenyi J. Association between obesity, quality of life, physical activity and health service utilization in primary care patients with osteoarthritis. *Int J Behav Nutr Phys Act* 2008;5:4.
- (12) Saag KG, Doebbeling BN, Rohrer JE, Kolluri S, Mitchell TA, Wallace RB. Arthritis health service utilization among the elderly: the role of urban-rural residence and other utilization factors. *Arthritis Care Res* 1998;11:177-85.
- (13) Thorstensson CA, Gooberman-Hill R, Adamson J, Williams S, Dieppe P. Help-seeking behaviour among people living with chronic hip or knee pain in the community. *BMC Musculoskelet Disord* 2009;10:153.
- (14) Snijders GF, den Broeder AA, van Riel PL, Straten VH, de Man FH, van den Hoogen FH, et al. Evidence-based tailored conservative treatment of knee and hip osteoarthritis: between knowing and doing. *Scand J Rheumatol* 2011;40:225-31.
- (15) Hunter DJ, Neogi T, Hochberg MC. Quality of osteoarthritis management and the need for reform in the US. *Arthritis Care Res (Hoboken)* 2011;63:31-8.
- (16) Smink AJ, van den Ende CH, Vliet Vlieland TP, Swierstra BA, Kortland JH, Bijlsma JW, et al. "Beating osteoARThritis": development of a stepped care strategy to optimize utilization

- and timing of non-surgical treatment modalities for patients with hip or knee osteoarthritis. *Clin Rheumatol* 2011;30:1623-9.
- (17) Linsell L, Dawson J, Zondervan K, Randall T, Rose P, Carr A, et al. Prospective study of elderly people comparing treatments following first primary care consultation for a symptomatic hip or knee. *Fam Pract* 2005;22:118-25.
 - (18) McHugh GA, Luker KA, Campbell M, Kay PR, Silman AJ. A longitudinal study exploring pain control, treatment and service provision for individuals with end-stage lower limb osteoarthritis. *Rheumatology (Oxford)* 2007;46:631-7.
 - (19) Hawker GA, Badley EM, Croxford R, Coyte PC, Glazier RH, Guan J, et al. A population-based nested case-control study of the costs of hip and knee replacement surgery. *Med Care* 2009;47:732-41.
 - (20) Andersen R, Newman JF. Societal and individual determinants of medical care utilization in the United States. *Milbank Mem Fund Q Health Soc* 1973;51:95-124.
 - (21) Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J Health Soc Behav* 1995;36:1-10.
 - (22) Jacobi CE, Triemstra M, Rupp I, Dinant HJ, Van Den Bos GA. Health care utilization among rheumatoid arthritis patients referred to a rheumatology center: unequal needs, unequal care? *Arthritis Rheum* 2001;45:324-30.
 - (23) Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. *Ann Rheum Dis* 2009;68:1413-9.
 - (24) Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;29:1039-49.
 - (25) Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991;34:505-14.
 - (26) Holla JF, Steultjens MP, van der Leeden M, Roorda LD, Bierma-Zeinstra SM, den Broeder AA, et al. Determinants of range of joint motion in patients with early symptomatic osteoarthritis of the hip and/or knee: an exploratory study in the CHECK cohort. *Osteoarthritis Cartilage* 2011;19:411-9.
 - (27) Swinkels H. Het meten van medische consumptie met behulp van enquêtes. In: *Het meten van determinanten van gezondheid: een overzicht van beschikbare meetinstrumenten*. 332-357. 1995. Assen: Van Gorcum. Ref Type: Report.
 - (28) Verstappen SM, Boonen A, Verkleij H, Bijlsma JW, Buskens E, Jacobs JW. Productivity costs among patients with rheumatoid arthritis: the influence of methods and sources to value loss of productivity. *Ann Rheum Dis* 2005;64:1754-60.
 - (29) Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs* 2005;14:798-804.
 - (30) Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833-40.

- (31) Veenhof C, Bijlsma JW, van den Ende CH, van Dijk GM, Pisters MF, Dekker J. Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature. *Arthritis Rheum* 2006;55:480-92.
- (32) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
- (33) Perrot S, Poiraudou S, Kabir M, Bertin P, Sichere P, Serrie A, et al. Active or passive pain coping strategies in hip and knee osteoarthritis? Results of a national survey of 4,719 patients in a primary care setting. *Arthritis Rheum* 2008;59:1555-62.
- (34) Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. *Br J Rheumatol* 1997;36:786-93.
- (35) Fink P, Jensen J, Borgquist L, Brevik JI, Dalgard OS, Sandager I, et al. Psychiatric morbidity in primary public health care: a Nordic multicentre investigation. Part I: method and prevalence of psychiatric morbidity. *Acta Psychiatr Scand* 1995;92:409-18.
- (36) Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthritis. *Ann Rheum Dis* 1957;16:494-502.
- (37) Steyerberg EW. *Clinical Prediction Models: A practical approach to development, validation, and updating.*, New York, Springer Sciences + Business Media, LLC 2009.
- (38) Twisk JWR. *Applied Longitudinal Data Analysis for Epidemiology: A practical guide.*, New York, Cambridge University Press 2003.
- (39) Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006;59:1092-101.
- (40) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
- (41) Royston P. Multiple imputation of missing values: Update of ice. *Stata Journal* 2005;5:527-36.
- (42) Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;9:57.
- (43) Albert SM, Musa D, Kwok CK, Hanlon JT, Silverman M. Self-care and professionally guided care in osteoarthritis: racial differences in a population-based sample. *J Aging Health* 2008;20:198-216.
- (44) Wandel S, Juni P, Tendal B, Nuesch E, Villiger PM, Welton NJ, et al. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010;341:c4675.
- (45) Cronan TA, Walen HR. Accuracy of self-reported healthcare use in patients with osteoarthritis. *J Rheumatol* 2002;29:2181-4.

Part II

Tailoring care for people with
advanced osteoarthritis

7

A multi-disciplinary and multidimensional
intervention for patients with hand
osteoarthritis.

ABSTRACT

Background. Although several guidelines recommend that treatment programs in patients with hand osteoarthritis (hand OA) should be both multidisciplinary and multidimensional, currently no suchlike treatment program for hand OA has been described. Therefore the aim of this study was to systematically develop a multidisciplinary and multidimensional non-pharmacological treatment program and to give a detailed description on the content of this treatment program.

Development of the treatment program. The program was developed in phases. In a pre-clinical theoretical phase disease specific problems, current evidence and, the influence of patients' characteristics on the benefit of interventions were explored. In a modelling phase the treatment program was designed.

Treatment program. the program contains an individual intake, four weekly nurse and occupational therapist led group sessions and, a booster session after six months. Treatment components of the program are self-management, daily home exercises to enhance joint mobility and grip strength and, education about ergonomic principles. The treatment program is tailored to the needs of individual patients with hand OA.

Conclusions. A non-pharmacological multidisciplinary and multidimensional treatment program for patients with hand OA was developed. Further research is necessary to investigate the effectiveness of this treatment program. Currently, the program is being evaluated in an ongoing randomised clinical trial.

Stukstette MJPM, Hoogeboom TJ, de Ruiter R, Koelmans P, Veerman E, den Broeder AA, Cats H, Bijlsma JW, Dekker J, van den Ende CHM.

A multi-disciplinary and multidimensional intervention for patients with hand osteoarthritis.

Clinical Rehabilitation 2012;26,99.

BACKGROUND

Osteoarthritis (OA) is the most common joint disease frequently affecting the joints of the hands (1). In elderly people (> 70 years) the prevalence of symptomatic hand OA is estimated to be 26% for women and 13% for men (2). Symptomatic hand OA has a great impact on daily functioning (3) and health related quality of life (4). Patients suffer from pain and joint stiffness and grip strength is reduced with more than 40% compared to healthy persons (3). Also, on average, hand mobility is restricted in all digits, in particular digits I, II and III (5). Furthermore, patients with hand OA suffer from fatigue, limitations in activities and participation (2,3,5-12).

Pharmacological treatment recommended for hand OA includes paracetamol, nonsteroidal anti-inflammatory drugs (NSAIDs), and symptomatic slow acting drugs for osteoarthritis (SYSADOAs) (2). Pharmacological treatments are symptomatic, as none of these treatments have demonstrated to retard the osteoarthritis features. Therefore, persons with hand OA are commonly referred to non-pharmacological treatment modalities to improve functional performance (13). A combination of pharmacological and non pharmacological treatment modalities is also strongly advised in the EULAR recommendations for treatment of patients with hand OA (14). However, no evidence and no guidelines are available about the optimal content and mix of non-pharmacological components (14).

In a cross sectional study on activity limitations and participation restrictions in patients with hand OA it was concluded that activity limitations are closely associated with hand impairments, while performance and satisfaction are more influenced by personal factors (3). As a consequence Kjekken et al. recommend that rehabilitation programmes should be provided by different health professionals (multidisciplinary) and focus on multiple dimensions (activities, participation and personal factors) of the International Classification of Functioning (ICF) (3). A rehabilitation program that consists of a number of components that may act both independently and inter-independently (such as in both multidisciplinary and multidimensional interventions) is called a complex intervention (15). By our knowledge, up till now no suchlike program for hand OA has been developed. This is confirmed in recent systematic reviews in which treatment modalities for hand OA were evaluated (16-19). None of the interventions evaluated were both multidisciplinary and multidimensional.

In the development of complex interventions some recommendations must be taken into account. First, several authors underline that a systematic approach to the development of a complex intervention leads to more effective interventions (20,21). Furthermore, Haynes et al. 1996 indicate that the development of an evidence based treatment program should synthesise information from three key sources with equal weighting, namely expert opinions, patient perspectives, and research evidence (22). A framework for the

development and evaluation of complex interventions to improve patients' health, has been proposed by the Medical Research Council (MRC) (15,23).

The aim of the current manuscript is twofold. The first aim is to describe how we systematically developed a complex non-pharmacological treatment program for patients with hand OA, following a stepwise approach, with reference to the MRC framework (15,23) and the recommendations of Haynes et al. to incorporate expert opinions, patient perspectives, and research evidence (22). A problem in implementing research evidence in clinical practice is a lack of treatment details in the description of trials and reviews (24). Therefore the second aim of this manuscript is to give a detailed description on the content of the treatment program developed. The evaluation of the effectiveness of this multidisciplinary treatment program is in progress (Dutch Trial Register trial number NTR1191). This paper does not provide the results of the evaluation of the intervention.

Development of the treatment program

According to the framework for the development and evaluation of complex interventions of the MRC we distinguished three phases in the development and evaluation of the treatment program (15,23): 1. a pre-clinical theoretical phase to identify disease specific problems; to explore the current evidence for treatment modalities and; to explore the influence of patients' characteristics on the benefit of intervention programs. 2. a modelling phase to formulate treatment goals, to weigh and prioritise the relevance of potential components and to decide on the set-up of the treatment program (content, number and frequency of sessions, duration of sessions), and 3. an exploratory phase to investigate the effectiveness of the treatment program. With respect to Haynes (22) we incorporated expert opinions, patient perspectives, and research evidence in each step. The steps we followed during the different phases are described in more detail below.

In the pre-clinical and theoretical phase specific problems due to hand OA and the evidence base for non-pharmacological treatment were explored. Disease specific problems due to hand OA were identified by a literature search in MEDLINE[®] from 1990 to October 2007. Because research into the effectiveness of non-pharmacological treatment in patients with hand OA is scarce we extended our search to studies on the effectiveness of non-pharmacological treatment in patients with hip or knee OA. We also analysed the results of interviews of 50 patients with hand OA referred to the multidisciplinary hand clinic of the Sint Maartenskliniek, a specialised rheumatology hospital in the Netherlands. Interviews were conducted according to the Canadian Occupational Performance Measure (COPM) (25) as part of routine care. The COPM is a semi-structured interview specially designed to identify and prioritise patient specific problematic activities (25). In an expert panel meeting the results of the interviews and the findings from the literature were discussed and supplemented in case of missing information. The expert panel consisted of two

rheumatologists, an occupational therapist, a specialized nurse, a physiotherapist, and two physiotherapy researchers of the Sint Maartenskliniek.

According to Newman et al. 2004, knowledge of the influence of patients' characteristics on treatment outcomes is needed to improve the benefit of intervention programs (26). Determinants of pain and limitations of activities were identified on the basis of a broad literature search concerning hand, hip and knee OA.

The modeling phase started with a meeting of the expert panel to formulate treatment goals. Assumptions used in the formulation of treatment goals were: 1. treatment goals should match specific problems due to hand OA as formulated in the pre-clinical and theoretical phase; 2. treatment goals should not conflict; 3. treatment goals should be realistic and feasible in a rheumatology outpatient centre.

Then, components of the treatment program matching treatment goals were selected utilising information of the pre-clinical theoretical phase. Identified components and treatment elements were discussed with patients with hand OA.

After formulating treatment goals and selection of components the format of the treatment program was designed by the expert panel. The expert panel decided for each treatment component on the built-up and time and number of sessions needed to attain treatment goals. A slide presentation for each treatment session, a therapist manual, and a patient manual were developed (available in Dutch, send request to the author).

In a pilot with 7 patients referred for multidisciplinary treatment due to hand OA, patients and therapists were asked to comment on the treatment program and materials used. Based on this pilot the protocol was adapted and refined.

Content and justification of the treatment program

Treatment goals

Hand OA has consequences on the levels of body functions and structures (pain, joint stiffness, fatigue, reduced muscle strength and joint mobility (2,3,5-12,27-32), activities (for example opening jars and bottles, doing and undoing buttons, riding a bike) and participation (self-care, leisure activities, commuting, and job participation) (2,3,5,32). After synthesizing the information from the literature with the results of our inventory of problematic activities the following goals covering all these levels were formulated:

1. The patient has control over pain and joint stiffness;
2. The patient has control over fatigue;
3. The patient is able to maintain / improve muscle strength and joint mobility;
4. The patient can anticipate difficulties in limitations in activities;
5. The patient can anticipate difficulties on the participation level;

Table 1. Final treatment program

Session	Duration (Hours)	Clinicians	Treatment components
Intake	1.5	OT	<ul style="list-style-type: none">• Screening• Formulating individual treatment goals
1	2.5	OT & SN	<ul style="list-style-type: none">• Self management to enhance self efficacy to cope with pain, limitations in activities and participation• Ergonomic principles• Daily home exercises to improve strength and hand mobility
2	3	OT & SN	<ul style="list-style-type: none">• Self management to enhance self efficacy to cope with pain, fatigue, limitations in activities and participation• Ergonomic principles• Daily home exercises to improve strength and hand mobility
3	2.5	OT & SN	<ul style="list-style-type: none">• Self management to enhance self efficacy to cope with pain, fatigue, limitations in activities and participation• Referral for splints if necessary• Daily home exercises to improve strength and hand mobility• Ergonomic principles
4	2.5	OT & SN	<ul style="list-style-type: none">• Self management to enhance self efficacy to cope with pain, fatigue, limitations in activities and participation• Daily home exercises to improve strength and hand mobility

Abbreviations: OT = Occupational Therapist, SN = Specialized Nurse

(Continuation Table 1).

Content of the session
<ul style="list-style-type: none"> • Screenings questionnaire • Semi structured interview (COPM)
<ul style="list-style-type: none"> • Discussion about common problems due to hand OA • Education about OA and treatment options (medication, splints, exercise) • Ergonomic principles • Practicing exercises to increase grip strength and joint mobility • Instruction to fill in a diary (activities, pain, fatigue and joy)
<ul style="list-style-type: none"> • Discussion about the relation of daily activities versus pain, fatigue joy and limitations in activities by means of discussing diaries • Education and discussion about strategies to cope with pain, fatigue and limitations in activities (balance between rest and physical activity, individual use of medication, use of compensatory strategies in daily activities, use of aids, use of splints) • Demonstration of aids and practicing with aids and ergonomics • Practicing exercises to increase grip strength and joint mobility • Formulating an individual action plan to achieve individual treatment goals
<ul style="list-style-type: none"> • Discussion on the progress of the action plan to achieve individual treatment goals • Discussion about problems experienced in daily life and possible solutions by means of the diaries • Discussion and reflection on strategies to cope with pain, fatigue and limitations in activities (balance between rest and physical activity, individual use of medication, use of compensatory strategies in daily activities, use of aids) • Education and discussion about splints (referral for splints if necessary); • Practicing exercises to increase grip strength and joint mobility • Practicing with aids and ergonomics in a kitchen
<ul style="list-style-type: none"> • Discussion about the relation of daily activities versus pain, fatigue joy and limitations in activities by means of discussing diaries • Discussion about strategies to cope with pain, fatigue and limitations in activities (balance between rest and physical activity, individual use of medication, use of compensatory strategies in daily activities, use of aids, use of splints) • Discussion on the progress of individual treatment goals • Discussion about problems experienced in daily life and possible solutions by means of the diaries • Practicing exercises to increase grip strength and joint mobility • Evaluation if patients' personal goals were established and evaluation of the treatment program
Abbreviations: OT = Occupational Therapist, SN = Specialized Nurse

Treatment program

Individual intake. Only certain subgroups of patients participating in non-pharmacological interventions successfully reach and maintain target physical performance levels on the short term. One of the most frequently mentioned reasons for these phenomena is lack of patients' motivation (29). Patient motivation is closely linked to "goal ownership", in other words, the degree to which patients consider the target health behaviour as their own self-chosen personal goal (33-35). Therefore, setting individual treatment goals is an important component of the treatment program in order to enhance the effectiveness of treatment (36-38).

Experienced limitations in activities and participation may vary among patients and are also influenced by the environment in which patients interact. Therefore, it was decided that patients have to be facilitated in formulating their own personal treatment goals in an individual intake session with a duration of one and a half hour. During the intake the COPM interview is conducted. The COPM is unique in that it obtains patient specific information in a structured way and helps patients in prioritizing their own specific problems and helps patients formulating their own treatment goals (25). It is a very helpful tool to bring the theory of a client centred approach into practice and provides patient specific information which cannot be obtained with other instruments that measure health (39). On the basis of the COPM, treatment goals are preferably set following the SMART principle, that is: goals should be specific, measurable, jointly set, realistic and time based.

Multidisciplinary group based sessions. Given the broad range of activities and functions possibly affected by hand OA the intake is followed by four weekly multidisciplinary sessions led by a hand therapist and a specialised rheumatology nurse. To ensure both group learning and sufficient individual guidance the treatment program is delivered in a group setting, with a minimum of 3 and a maximum of 9 participants per group. Six months after finishing the treatment program a booster session is planned to reinforce adherence. Below the treatment components are described. The treatment program is summarised per session in Table 1.

Treatment component 1 – Self-management to enhance self efficacy to cope with pain, fatigue, limitations in activities, and participation(40-46). Self-management interventions result in a reduction in pain, fatigue, and health distress, as well as an increase in self-efficacy and physical activity. Self-management is the individual's ability to manage the symptoms, treatment, physical and psychosocial consequences and life style changes inherent in living with a chronic disease (47). Patients with arthritis view self management interventions as a way to bring order into their lives, to recognise boundaries plan, pace and prioritise and, to cope with change in self-identity (48). Effective self management interventions are problem focused, action oriented and utilize patient generated care plans (49).

To enhance patients self management the following topics are part of our program:

- Education and discussion about hand OA and its treatment options: medication, non-pharmacological treatment options, life style changes, exercises to improve joint mobility and muscle strength, ergonomic principles, aids, splints;
- Discussion and reflection upon the influence of the nature and intensity of daily activities on pain, fatigue, and limitations in activities by means of discussing activity diaries;
- Discussion and reflection on strategies to cope with pain, fatigue and limitations in activities (balance between rest and physical activity, individual use of medication, use of compensatory strategies in daily activities, use of aids, use of splints);
- Formulating an individual action plan to achieve individual goals.

Treatment component 2 - Daily home exercises to improve strength and hand mobility: Patients with hand OA suffer from reduced grip strength and reduced joint mobility in particular in digits I, II and III and the first carpometacarpal joint (CMC I-joint). Joint mobility and muscle strength, in particular grip and pinch strength are related with limitations in daily activities and pain. Studies have shown the effectiveness of training muscle strength on global hand function (50) and the effectiveness of training muscle strength and joint mobility on levels of pain (51). Because improvement of joint mobility and muscle strength could lead to reduced limitations in activities and pain an exercise program was incorporated in the treatment program. In Table 2 this exercise program is described in more detail.

To integrate exercises in daily life activities, patients are encouraged to record on a daily basis the performed exercises and their experiences. Our experience is that exercise diaries are of great value to give health professionals important information on facilitators and barriers to implement exercises in daily life.

Treatment component 3 - Ergonomic principles: in patients with rheumatoid arthritis the joint protection program has shown to have positive effects on hand pain, joint stiffness, and limitations in activities (52). In treating patients with hand OA a combination of joint protection and exercise therapy was shown to be effective (50). The following elements of the joint protection program were considered relevant and adopted in our treatment program:

- The use of large instead of smaller joints (for example by carrying a bag with the handles on the forearm)
- The use of more joints at a time (e.g. preferably holding a mug with two hands).
- The use of practical aids to unburden affected joints (for example using a rubberised jar opener, adhesive foam or custom handles, pen or pencil holding devices, or house/car key adaptations).

Table 2. Exercises to enhance joint mobility and muscle strength.

Exercise number	Name of exercise	Patient instruction
1	Mobility of the wrist and finger joints	Put your arms outstretched on a table with your palms together. Move your wrists as far as possible toward your body by bending your elbows. keep your hands together and keep your fingers extended.
2	Mobility of the wrist and finger joints	Take a stick loosen in your hands. Create a wringing movement with your wrists. Make the movements with your wrists as large as possible. Please note that you will not pinch in the stick.
3	Mobility of the finger joints	Move your fingertips as far as possible to the base of your fingers (make a small fist). Stretch your fingers and make a completely normal fist (make a major fist), stretch your fingers completely and start over again.
4	Mobility of the thumb base joint	Put your hand outstretched with your palm upwards on the table. Move the tip of your thumb as far as possible to the base of your finger. Then move your thumb stretched as far as possible to the table. Make sure the angle between your forefinger and your thumb is as big as possible.
5	Strength of the intrinsic muscles	Take a piece of clay in your hands and knead the clay.
6	Strength of the intrinsic muscles	Take a ball of clay between your extended thumb and fingers. Press the clay with your fingers extended so far as possible to your thumb. The lower hand acts as a pallet and to turn the ball of clay. You can do this exercise until you notice that your muscles begin to feel tired
7	Strength of the extensor muscles	Take a piece of clay in your hands. Pull the clay with your fingertips as far as possible up in your palm and start over again. You can do this exercise until you notice that your muscles begin to feel tired.
8	Coordination of the fingers	Take a ball between your thumb and forefinger. Move the ball as far as possible over the outstretched thumb to the back and away over the outstretched thumb to the front. Repeat the exercise with all your fingers. You can do these exercises until you feel your muscles begin to feel tired.

Photographs of all exercises are available at the website of clinical rehabilitation

To facilitate patients in finding a technique or aid that fits his or her needs, we experienced that it is important to instruct patients in the use of ergonomic principles by providing them the opportunity to practice various techniques and to experiment with practical aids in more than one session.

Treatment component 4 - Referral for a splint if considered necessary by a hand therapist: Several studies investigated the effectiveness of splints in patients with hand OA. Although the quality of most of these studies is weak due to a small study sample or an inappropriate study design, there are indications that the use of splints leads to reduced pain and limitations of activities (14,16-19). Patients are referred for a splint if considered necessary.

Built up of the treatment program. Each component is repeated and built up during several sessions. For example: after formulating and prioritising individual treatment goals at the intake, during the group sessions patients make an individual action plan which is evaluated by means of discussing diaries during each session. Another example: exercises are part of every session; in the first session patients are instructed to determine their basic level while in the last two session ways to integrate exercises in daily life are emphasised. For treatment components with a theoretical and a practical part, it was decided that the theoretical part has to be discussed before practicing is started. Ergonomic principles for example in the first session were discussed and practiced in a kitchen in the second and third session.

Table 3. Take home messages.

-
- OA: it cannot be cured, but there is something to do to deal with it
 - Know which medication you swallow and how and when to swallow it
 - Ergonomic principles can help you act easier in daily life
 - Be smart and act easier
 - Normal load does not lead to additional joint damage
 - Exercise and normal use of hands does not lead to additional joint damage
 - Maintaining / improving muscle strength helps you to continue to perform activities
 - Pain is not always a reason to worry
 - Know your limit and dare to set your limit
 - Communicate with people in your environment
 - You are experienced with your disease, find out what helps you
 - Exercises help you to reduce joint stiffness
-

Abbreviations: OA = Osteoarthritis

During all sessions discussion among patients is encouraged to utilise patients' expert opinions. Furthermore, each session includes clearly stated take home messages; such as "normal use of hands and exercises will not damage your joints". In Table 3 the take home messages are presented.

DISCUSSION

The aim of this study was to systematically develop a non-pharmacological multidisciplinary and multidimensional treatment program for patients with hand OA. This resulted in a reproducible treatment program consisting of self management, daily home exercises and education about ergonomic principles, tailored to the needs of the individual patient with hand OA.

In all steps of the development of the treatment program we searched for evidence from the literature, asked patients opinions where possible and in case of lacking information also expert opinions were gathered. The involvement of patients was especially important in the identification of patient specific problems. Patients' specific problems depend on the environment in which they interact and therefore are influenced by cross-cultural differences. For example, in the Netherlands the main transport of many people is a bicycle. Problems experienced in holding the steer of a bicycle and using the handbrakes of a bike are frequently mentioned problems in patients with hand OA. For that reason, possible limitations in biking were incorporated in the treatment program and patient materials.

A possible limitation in the development process was that all members of the expert panel involved in developing the multidisciplinary treatment program were employed in the same rheumatology outpatient clinic. As a result, possibly the views and opinions of the expert panel do not fully represent those of other health professionals nationally and internationally. On the other hand, the involvement of multiple disciplines within our setting helped us to create acceptance and support for the implementation of the treatment program. Furthermore, by including multiple disciplines of our hospital in the expert panel we ensured a match with the opportunities and facilities within our hospital. It is possible that other settings do not have the same facilities (for example a kitchen to practice with ergonomic principles and aids) and opportunities (for example funding structures) as we have. Therefore, it is likely that the program is not feasible in its original extent in all other settings.

It has been suggested to use a causal modelling approach or a method of concept mapping to define the target group and the desirable health outcome (53,54). However, those approaches do not provide practical methods to translate evidence into practice. Therefore, we utilised a more empirical perspective. Our starting points were the nature of problems of patients with hand OA seeking medical care and the recommendations outlined in guidelines about non-pharmacological treatment for this category of patients. We elaborated on the MRC framework resulting in a comprehensive treatment program consisting of components proven to be effective in OA and tailored to the needs of patients with hand OA. Our study might be a valuable illustration of a feasible and practical method to integrate evidence, expertise, and clinical practice into treatment programs for multidisciplinary teams.

To investigate the effectiveness of our newly developed treatment program we started a multicentre randomised single blind controlled clinical trial in 150 patients with hand OA according to the American College of Rheumatology Classification criteria for Hand OA(55). The primary outcome measures after three months are the OMERACT OARSI responder criteria (55,56) and the Australian Canadian Osteoarthritis Hand Index (31,57).

In the development process of the treatment program patients indicated that they would prefer a booster session to discuss how strategies are implemented in daily life and how patients could further improve. In patients with hip and knee OA booster sessions seem to

maintain the effects of exercise therapy on the long term (58). Given the lack of evidence about the effectiveness of booster sessions in patients with hand OA, we will also investigate the effectiveness of a booster session six months after a multidisciplinary treatment program.

In conclusion: a non-pharmacological multidisciplinary and multidimensional treatment program for patients with hand OA was developed following a systematic approach, utilizing the current evidence and patient and expert opinions. Further research is now performed to investigate the effectiveness of this treatment program.

REFERENCE LIST

- (1) Creamer P, Hochberg MC. Osteoarthritis. *Lancet* 1997 08/16//350(9076):503-508.
- (2) Zhang Y, Niu J, Kelly-Hayes M, Chaisson CE, Aliabadi P, Felson DT. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly: The Framingham Study. *Am J Epidemiol* 2002 12/01//156(11):1021-1027.
- (3) Kjekouk I, Dagfinrud H, Slatkowsky-Christensen B, Mowinckel P, Uhlig T, Kvien TK, et al. Activity limitations and participation restrictions in women with hand osteoarthritis: patients' descriptions and associations between dimensions of functioning. *Ann Rheum Dis* 2005 11//64(11):1633-1638.
- (4) Slatkowsky-Christensen B, Mowinckel P, Loge JH, Kvien TK. Health-related quality of life in women with symptomatic hand osteoarthritis: a comparison with rheumatoid arthritis patients, healthy controls, and normative data. *Arthritis Rheum* 2007 12/15//57(8):1404-1409.
- (5) Spacek E, Poiraudou S, Fayad F, Lefevre-Colau MM, Beaudreuil J, Rannou F, et al. Disability induced by hand osteoarthritis: are patients with more symptoms at digits 2-5 interphalangeal joints different from those with more symptoms at the base of the thumb? *Osteoarthritis Cartilage* 2004 05//12(5):366-373.
- (6) Bergman MJ, Shahouri SS, Shaver TS, Anderson JD, Weidensaul DN, Busch RE, et al. Is fatigue an inflammatory variable in rheumatoid arthritis (RA)? Analyses of fatigue in RA, osteoarthritis, and fibromyalgia. *J Rheumatol* 2009 12//36(12):2788-2794.
- (7) Murphy SL, Smith DM, Clauw DJ, Alexander NB. The impact of momentary pain and fatigue on physical activity in women with osteoarthritis. *Arthritis Rheum* 2008 06/15//59(6):849-856.
- (8) Parrish BP, Zautra AJ, Davis MC. The role of positive and negative interpersonal events on daily fatigue in women with fibromyalgia, rheumatoid arthritis, and osteoarthritis. *Health Psychol* 2008 11//27(6):694-702.
- (9) Power JD, Badley EM, French MR, Wall AJ, Hawker GA. Fatigue in osteoarthritis: a qualitative study. *BMC Musculoskelet Disord* 2008;9:63.
- (10) Stebbings S, Herbison P, Doyle TC, Trehan GJ, Highton J. A comparison of fatigue correlates in rheumatoid arthritis and osteoarthritis: disparity in associations with disability, anxiety and sleep disturbance. *Rheumatology (Oxford)* 2010 02//49(2):361-367.
- (11) Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. *Rheumatology (Oxford)* 1999 04//38(4):355-361.
- (12) Zautra AJ, Fasman R, Parish BP, Davis MC. Daily fatigue in women with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Pain* 2007 03//128(1-2):128-135.
- (13) Brandt KD. The importance of nonpharmacologic approaches in management of osteoarthritis. *Am J Med* 1998 07/27//105(1B):39S-44S.
- (14) Zhang W, Doherty M, Leeb BF, Alekseeva L, Arden NK, Bijlsma JW, et al. EULAR evidence based recommendations for the management of hand osteoarthritis: report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007 03//66(3):377-388.

- (15) Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000 09/16/;321(7262):694-696.
- (16) Egan MY, Brousseau L. Splinting for osteoarthritis of the carpometacarpal joint: a review of the evidence. *Am J Occup Ther* 2007 01//;61(1):70-78.
- (17) Mahendira D, Towheed TE. Systematic review of non-surgical therapies for osteoarthritis of the hand: an update. *Osteoarthritis Cartilage* 2009 10//;17(10):1263-1268.
- (18) Moe RH, Kjekken I, Uhlig T, Hagen KB. There is inadequate evidence to determine the effectiveness of nonpharmacological and nonsurgical interventions for hand osteoarthritis: an overview of high-quality systematic reviews. *Phys Ther* 2009 12//;89(12):1363-1370.
- (19) Towheed TE. Systematic review of therapies for osteoarthritis of the hand. *Osteoarthritis Cartilage* 2005 06//;13(6):455-462.
- (20) Green LWK, M.W. Health promotion planning: an Educational and Ecological Approach. 3rd ed. Mountain View, California: Mayfield Publishing Company; 1999.
- (21) Kok G, van den Borne B, Mullen PD. Effectiveness of health education and health promotion: meta-analyses of effect studies and determinants of effectiveness. *Patient Educ Couns* 1997 Jan;30(1):19-27.
- (22) Haynes RB, Sackett DL, Gray JM, Cook DJ, Guyatt GH. Transferring evidence from research into practice: 1. The role of clinical care research evidence in clinical decisions. *ACP J Club* 1996 11//;125(3):A14-A16.
- (23) Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655.
- (24) Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008 06/28/;336(7659):1472-1474.
- (25) Law M, Baptiste S, McColl M, Opzoomer A, Polatajko H, Pollock N. The Canadian occupational performance measure: an outcome measure for occupational therapy. *Can J Occup Ther* 1990 04//;57(2):82-87.
- (26) Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. *Lancet* 2004 10/23/;364(9444):1523-1537.
- (27) Bellamy N, Campbell J, Haraoui B, Gerez-Simon E, Buchbinder R, Hobby K, et al. Clinimetric properties of the AUSCAN Osteoarthritis Hand Index: an evaluation of reliability, validity and responsiveness. *Osteoarthritis Cartilage* 2002 11//;10(11):863-869.
- (28) Allen KD, Coffman CJ, Golightly YM, Stechuchak KM, Keefe FJ. Daily pain variations among patients with hand, hip, and knee osteoarthritis. *Osteoarthritis Cartilage* 2009 10//;17(10):1275-1282.
- (29) Bijsterbosch J, Visser W, Kroon HM, Stamm T, Meulenbelt I, Huizinga TW, et al. Thumb base involvement in symptomatic hand osteoarthritis is associated with more pain and functional disability. *Ann Rheum Dis* 2010 03//;69(3):585-587.
- (30) Marshall M, van der Windt D, Nicholls E, Myers H, Hay E, Dziedziec K. Radiographic hand osteoarthritis: patterns and associations with hand pain and function in a community-dwelling sample. *Osteoarthritis Cartilage* 2009 11//;17(11):1440-1447.
- (31) Bellamy N, Sothorn RB, Campbell J, Buchanan WW. Rhythmic variations in pain, stiffness, and manual dexterity in hand osteoarthritis. *Ann Rheum Dis* 2002 12//;61(12):1075-1080.

- (32) Slatkowsky-Christensen B, Haugen I, Kvien TK. Distribution of joint involvement in women with hand osteoarthritis and associations between joint counts and patient-reported outcome measures. *Ann Rheum Dis* 2010 01//;69(1):198-201.
- (33) Joint Commission on Accreditation of Healthcare Organizations. Standards. Chicago, Illinois; 1992.
- (34) Commission on Accreditation of Rehabilitation Facilities. Commission on Accreditation of Rehabilitation Facilities Standards Manual and Interpretive Guidelines for Organizations Serving People With Disabilities. Tucson, Arizona; 1994.
- (35) Northen JG, Rust DM, Nelson CE, Watts JH. Involvement of adult rehabilitation patients in setting occupational therapy goals. *Am J Occup Ther* 1995 03//;49(3):214-220.
- (36) Payton O, Nelson C. A preliminary study of patients' perceptions of certain aspects of their physical therapy experience. *Physiotherapy Theory and Practice* 1996;12(1):27-38.
- (37) Payton O, Ozer M, Nelson C. Patient Participation in Program Planning: A Manual for Therapists. 1990;6.
- (38) Randall KE, McEwen IR. Writing patient-centered functional goals. *Phys Ther* 2000 12//;80(12):1197-1203.
- (39) Dedding C, Cardol M, Eyssen IC, Dekker J, Beelen A. Validity of the Canadian Occupational Performance Measure: a client-centred outcome measurement. *Clin Rehabil* 2004 Sep;18(6):660-667.
- (40) Buszewicz M, Rait G, Griffin M, Nazareth I, Patel A, Atkinson A, et al. Self management of arthritis in primary care: randomised controlled trial. *BMJ* 2006 10/28//;333(7574):879.
- (41) Devos-Comby L, Cronan T, Roesch SC. Do exercise and self-management interventions benefit patients with osteoarthritis of the knee? A metaanalytic review. *J Rheumatol* 2006 04//;33(4):744-756.
- (42) Heuts PH, de BR, Drietaalar M, Aretz K, Hopman-Rock M, Bastiaenen CH, et al. Self-management in osteoarthritis of hip or knee: a randomized clinical trial in a primary healthcare setting. *J Rheumatol* 2005 03//;32(3):543-549.
- (43) Lorig KR, Ritter P, Stewart AL, Sobel DS, Brown BW, Jr., Bandura A, et al. Chronic disease self-management program: 2-year health status and health care utilization outcomes. *Med Care* 2001 11//;39(11):1217-1223.
- (44) Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. *Eff Clin Pract* 2001 11//;4(6):256-262.
- (45) Osborne RH, Wilson T, Lorig KR, McColl GJ. Does self-management lead to sustainable health benefits in people with arthritis? A 2-year transition study of 452 Australians. *J Rheumatol* 2007 05//;34(5):1112-1117.
- (46) Warsi A, LaValley MP, Wang PS, Avorn J, Solomon DH. Arthritis self-management education programs: a meta-analysis of the effect on pain and disability. *Arthritis Rheum* 2003 08//;48(8):2207-2213.
- (47) Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. *Patient Educ Couns* 2002 Oct -Nov;48(2):177-187.
- (48) Kralik D, Koch T, Price K, Howard N. Chronic illness self-management: taking action to create order. *J Clin Nurs* 2004 Feb;13(2):259-267.

- (49) Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7.
- (50) Stamm TA, Machold KP, Smolen JS, Fischer S, Redlich K, Graninger W, et al. Joint protection and home hand exercises improve hand function in patients with hand osteoarthritis: a randomized controlled trial. *Arthritis Rheum* 2002 02//;47(1):44-49.
- (51) Rogers MW, Wilder FV. The effects of strength training among persons with hand osteoarthritis: a two-year follow-up study. *J Hand Ther* 2007 07//;20(3):244-249.
- (52) Hammond A, Freeman K. One-year outcomes of a randomized controlled trial of an educational-behavioural joint protection programme for people with rheumatoid arthritis. *Rheumatology (Oxford)* 2001 09//;40(9):1044-1051.
- (53) Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From Theory to intervention: Mapping theoretically derived behavioural determinants to behaviour change techniques. *Applied Psychology: An International Review* 2008;57(4):660-680.
- (54) Hardeman W, Sutton S, Griffin S, Johnston M, White A, Wareham NJ, et al. A causal modelling approach to the development of theory-based behaviour change programmes for trial evaluation. *Health Educ Res* 2005 12//;20(6):676-687.
- (55) Pham T, van der HD, Lassere M, Altman RD, Anderson JJ, Bellamy N, et al. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. *J Rheumatol* 2003 07//;30(7):1648-1654.
- (56) Pham T, van der HD, Altman RD, Anderson JJ, Bellamy N, Hochberg M, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. *Osteoarthritis Cartilage* 2004 05//;12(5):389-399.
- (57) Bellamy N, Campbell J, Haraoui B, Buchbinder R, Hobby K, Roth JH, et al. Dimensionality and clinical importance of pain and disability in hand osteoarthritis: Development of the Australian/Canadian (AUSCAN) Osteoarthritis Hand Index. *Osteoarthritis Cartilage* 2002 11//;10(11):855-862.
- (58) Pisters MF, Veenhof C, van Meeteren NL, Ostelo RW, de Bakker DH, Schellevis FG, et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review. *Arthritis Rheum* 2007 10/15//;57(7):1245-1253.

8

Non-pharmacological care for patients with
generalized osteoarthritis: design of a
randomized clinical trial.

ABSTRACT

Background. Non-pharmacological treatment (NPT) is a useful treatment option in the management of hip or knee osteoarthritis. To our knowledge however, no studies have investigated the effect of NPT in patients with generalized osteoarthritis (GOA). The primary aim of this study is to compare the effectiveness of two currently existing health care programs with different intensity and mode of delivery on daily functioning in patients with GOA. The secondary objective is to compare the cost-effectiveness of both interventions.

Methods/Design. In this randomized, single blind, clinical trial with active controls, we aim to include 170 patients with GOA. The experimental intervention consists of six self-management group sessions provided by a multi-disciplinary team (occupational therapist, physiotherapist, dietician and specialized nurse). The active control group consists of two group sessions and four sessions by telephone, provided by a specialized nurse and physiotherapist. Both therapies last six weeks. Main study outcome is daily functioning during the first year after the treatment, assessed on the Health Assessment Questionnaire. Secondary outcomes are health related quality of life, specific complaints, fatigue, and costs. Illness cognitions, global perceived effect and self-efficacy, will also be assessed for a responder analysis. Outcome assessments are performed directly after the intervention, after 26 weeks and after 52 weeks.

Discussion. This article describes the design of a randomized, single blind, clinical trial with a one year follow up to compare the costs and effectiveness of two non-pharmacological interventions with different modes of delivery for patients with GOA.

Trial registration. Dutch Trial Register NTR2137

Hoozeboom TJ, Stukstette MJPM, de Bie RA, Cornelissen J,
den Broeder AA, van den Ende CHM.

*Non-pharmacological care for patients with generalized osteoarthritis:
design of a randomized clinical trial.*

BMC Musculoskeletal Disorders 2010;11,142.

BACKGROUND

Osteoarthritis (OA), the most prevalent rheumatic disorder affecting the musculoskeletal system, has a major impact on functioning and independence of the elderly. In general, one distinguishes between four subgroups of peripheral OA: 1. knee, 2. hip, 3. hand and 4. generalized OA (GOA) [1]. Clinically, 10–15% of adults over 60 have symptomatic knee OA [2-5] and symptomatic hip OA occurs in 1–4% of all adults [3]. Symptomatic hand OA occurs in 10–15% of the elderly [6] and GOA in 27% of the patients with hip or knee OA [7].

Non-pharmacological treatment (NPT) is considered to be important in the management of OA in order to reduce the impact of OA on pain and physical functioning [8]. Current OA research on NPT-options, focuses mainly on the hip and knee joint [9]. An abundance of research literature illustrates that NPT is a useful treatment option in the management of hip or knee OA [9]. The initial focus of NPT should lie on self-management and patient-driven treatments, rather than on passive therapies delivered by allied health professionals [8]. Provision of information and patient education about the objectives of treatment and the importance of changes in lifestyle, exercise, pacing of activities, weight reduction and other measures to unload damaged joints is supported by two meta-analyses [10,11] on the efficacy of non-pharmacological interventions in chronic diseases. Increasing the functional capacity [12-14] and encouraging the patient to undertake and maintain regular exercise [15] have also been found effective.

There is lack of evidence concerning the optimum mode of care delivery. The more traditional face-to-face contact is by far the most evaluated type of therapy delivery. However, telephone contact aimed at promoting self-care appears to be more cost-efficient [16] and has also been associated with improvements in joint pain [17,18] and physical function [18] for up to a year in patients with knee OA. Moreover, in a recent study by Eakon et al (2009) telephone counselling was suggested a feasible mode of delivering lifestyle interventions to patients with chronic conditions and demonstrated modest improvements in diet and physical activity [19].

To our knowledge, only one study investigated the effect of a non-pharmacological intervention in the management of GOA [20]. However, in this study all GOA-patients recently underwent major joint replacement, therefore the results of this study cannot be generalized to a population in which joint replacement is not (yet) an option. Taking into account a. the extensive body of literature on NPT for hip or knee OA, b. the substantial group of patients and c. the fact OA in multiple joints is more disabling than in one joint [21-23], it is remarkable that research on the efficacy of NPT options in GOA is hitherto neglected.

Considering the latter, we infer that the development and evaluation of a treatment programme is warranted. To do so, we installed an expert group consisting of a

physiotherapist, an occupational therapist, a specialized nurse, a rheumatologist and two researchers, all of whom have extensive experience with GOA patients. Consequently, the expert group systematically conceptualized a definition of GOA and a treatment programme tailored to the needs of patients with GOA and based on recommendations for the management of hip and knee OA [8,9] and on the clinical experience of the health care providers. This resulted in a best-evidence, multi-disciplinary treatment programme. Since there is no information about the optimal treatment intensity and mode of delivery, we decided to compare the effectiveness of a fully supervised multi-disciplinary program to an active control [24] (i.e. a telephone monitored program combined with two supervised contact moments). Due to the complex nature of GOA and the fact that guidelines for hip and knee OA recommend multiple NPT modalities, both interventions are multi-disciplinary [8]. We hypothesize that both programmes have beneficial effects on the patients' quality of life and ability to cope with their disease, however we expect the face-to-face programme to be superior with respect to daily functioning to the telephone programme.

The primary aim of this study is to compare the effectiveness of a supervised multi-disciplinary programme to an active control on daily functioning in patients with GOA during the first year after treatment. Secondary aims of the study are to investigate the short-term effects of interventions and to compare the cost-effectiveness of both interventions.

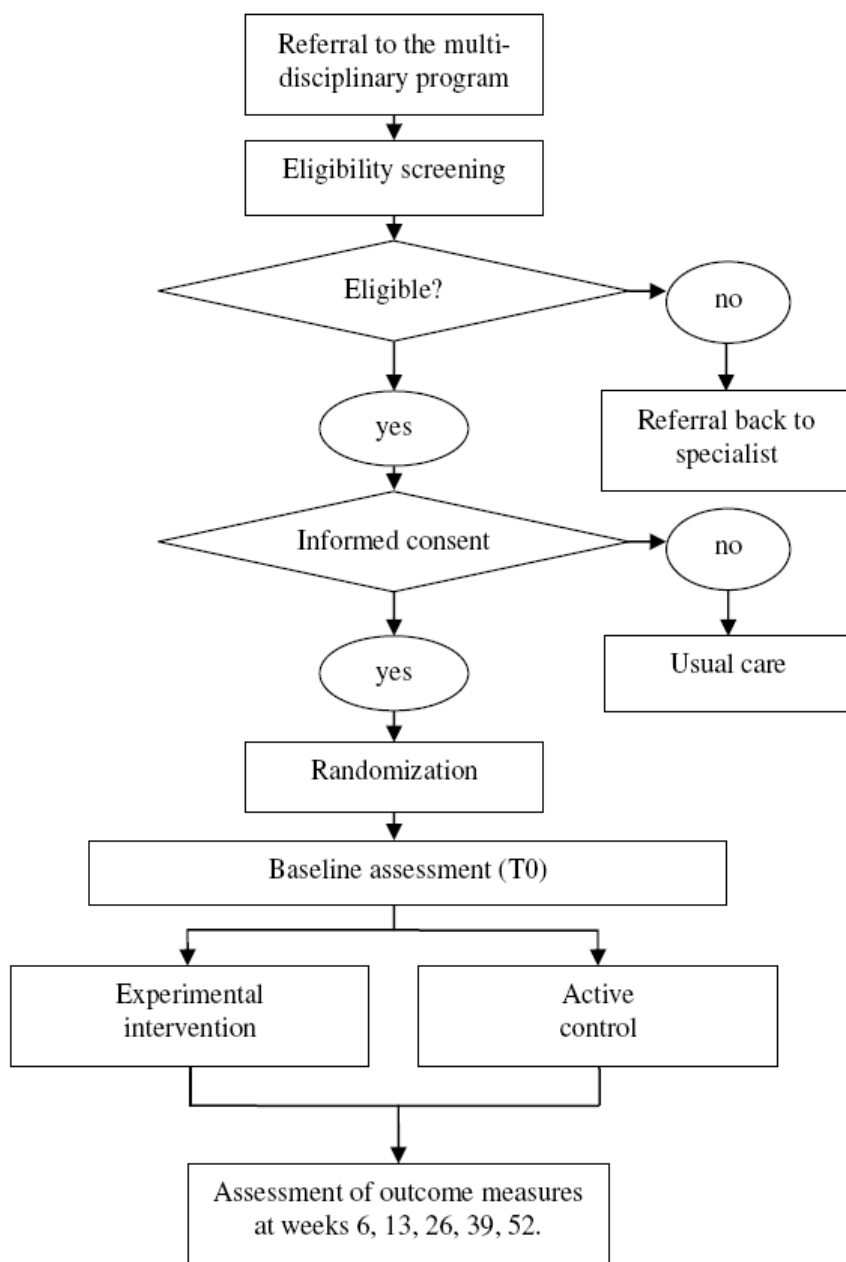
METHODS/DESIGN

A pragmatic randomized, single blind, clinical, superiority trial with active controls will be used to study the aforementioned aims. The study will be performed at the outpatient rheumatology departments of the Sint Maartenskliniek Hospitals in the cities of Woerden and Nijmegen in The Netherlands. Both centres have piloted the interventions and in both centres, rooms well equipped for group based treatments are available.

Patients - referred by their rheumatologist to the outpatient department for multi-disciplinary NPT - eligible for both the GOA health care program and the trial are informed about the trial. Subsequently, consenting patients are randomly allocated to one of the two groups and followed by questionnaires for a total of 52 weeks (figure 1).

The trial has been reviewed by the Institutional Review Board of the University Medical Centre Nijmegen (protocol number 2009/290) and they concluded that the study did not fall within the remit of the Medical Research Involving Human Subjects Act. So the study can be carried out (in the Netherlands) without an approval by an accredited ethical board.

Figure 1. An overview of the study design showing recruitment, assessments, treatment groups, follow-up and outcome assessments.



Eligibility Criteria

Men and women (≥ 18 years old) are eligible to enter the trial if they are diagnosed with GOA (see our definition in the following paragraph), motivated to alter their lifestyle (assessed by a standardized set of questions), willing to participate in a group and able to comply with the planned time schedule of both treatment conditions. Patients are excluded when they are 1. awaiting surgery, 2. already participated unsuccessfully in a self-management program, 3. are considered not to be able to participate in a group due to limited psychological functioning (on the basis of clinical judgment of a psychologist), 4. are illiterate, 5. are not capable of communicating in Dutch or 6. are incapable of coming to the hospital.

From the patients who were in principle eligible considering the in- and exclusion criteria but decided not to participate in the study, baseline demographics (ie. age and sex) will be gathered, to assess possible selection bias.

Definition of GOA

In the abovementioned inclusion criteria we mentioned that patients must be diagnosed with GOA. However, no uniform GOA definition is available in the literature. A pragmatic literature search elicited numerous definitions for GOA [7,25-33], mainly used in genetic studies and for the greater part based on patterns of distribution of joints with radiological changes.

In clinical practice the term GOA refers to the combination of clinical symptoms and radiographic changes in multiple joints which can be attributed to OA as obtaining a full picture of radiological changes in all joints is not feasible nor desirable in clinical practice. For the purpose of this project we formulated a pragmatic definition of GOA based upon literature findings and on the basis of consensus of several clinicians and health professionals with experience in patients with GOA. In our definition signs and symptoms are combined with radiological changes. In this project a patient is defined as having GOA if he or she complies with the following three conditions:

- a. experiencing complaints in three or more groups of joints, and;
- b. having at least two objective signs that indicate OA in at least two joints (objective signs indicating OA are malalignment, palpable osteophytes / nodules, crepitations over the full range of motion, and limited range of motion or radiographic signs including the presence of joint space narrowing and/or osteophytes, and;
- c. is limited in daily functioning (Health Assessment Questionnaire score [34] > 0.5).

Interventions

During a six-week treatment period, patients will receive one of the following two treatment programs:

- Interdisciplinary, group-based, self-management program (experimental intervention);

- Telephone-based, self-management program (active control).

Both pilot tested interventions were developed from a clinical and pragmatic perspective, meaning that both interventions had to be useful and feasible in clinical practice. This resulted in two, for patients and health care providers satisfactory, interventions. However, from a research perspective lack of contrast between both interventions might be observed. To further elaborate differences between groups, we depicted a specific overview of the content of the non-pharmacological treatment in both arms in Figure 2a and 2b, according to the recommendations of Perera *et al* 2007 [35].

For both interventions, manuals and standardized presentations were created. At baseline and every six months, all care providers assemble to assess and enhance the adherence with the treatment protocols.

Experimental intervention group

Patients (eight per group) allocated to the experimental intervention group attend six therapeutic group sessions and one group evaluation. During these six sessions patients aim to improve daily functioning by optimising their current lifestyle (i.e. physical activity and diet) and by enhancing self-efficacy to control the consequences of the disease in everyday life (i.e. activity pacing, pain management and daily functioning). To enhance patients' self-efficacy the 5As model of behaviour change counselling is used, which is an evidence-based approach appropriate for a broad range of different behaviours and health conditions. The 5As consists of: Assessing patient level of behaviour, beliefs and motivation; Advising the patient based upon personal health risks; Agreeing with the patient on a realistic set of goals; Assisting to anticipate barriers and develop a specific action plan; and Arranging follow-up support [36]. The following example might illustrate the use of the 5As model.

A participant wears a pedometer to elicit his/her physical activity level (A1). Together with the health care provider, the patient discusses the outcome (A2) and set a goal to increase the level of physical activity (A3). Both the health care provider and the patient must believe the goal is adequate (A3) and realistic (A4). Consequently, patient and therapist closely monitor the personal goals (A5). In addition to the self-management programme, patients are also enrolled in an exercise programme aimed to 1. improve the quality of movement and 2. implement the learned exercises in the home situation.

Figure 2. Graphical depiction of the two self-management interventions

Timeline	Experimental intervention	Active control
	General intake by PT and OT	
Eligibility assessment		
Baseline assessment		
Randomisation		
Week 1a	<div><div>a</div><div>b</div><div>a</div></div>	<div><div>a</div><div>b</div><div>c</div><div>d</div><div>e</div><div>f</div></div>
Week 1b	<div><div>c</div><div>d</div><div>a</div><div>b</div></div>	<div><div>f</div></div>
Week 2a	<div><div>e</div><div>a</div><div>c</div></div>	
Week 2b	<div><div>e</div><div>b</div><div>c</div><div>d</div></div>	<div><div>f</div></div>
Week 3	<div><div>f</div><div>c</div><div>d</div></div>	<div><div>c</div><div>d</div></div>
Week 4	<div><div>b</div><div>c</div><div>d</div></div>	<div><div>f</div></div>
Week 6	<div><div>e</div></div>	<div><div>e</div><div>f</div></div>

(continuation Figure 2).

Measurement of outcomes	
(a)	Group education on the health care program, diaries and expectations.
(b)	Group education on osteoarthritis, pain and medication.
a	General exercise program.
(c)	Group education about physical activity.
(d)	Group education on activity pacing.
b	Recreational activity.
c	Group based monitoring of personal goals.
(e)	Group education on food consumption.
d	Specific exercise program, based upon the PSK-scores.
(f)	Group education on acceptance and helplessness.
e	Evaluation and setting goals for the future.
f	Monitoring of personal goals via telephone by specialized nurse.
Abbreviations: PSK = Patient Specific Complaints, PT = physiotherapist, OT = occupational therapist	

Active control group

Patients enrolled in the active control group, attend two group sessions (eight patients per group) and are further monitored through four telephone contacts [37]. As per with the experimental intervention, the active control group aims to optimise the patients' current lifestyle (i.e. physical activity and diet) and to enhance the patients' self-efficacy to control the disease (i.e. activity pacing, pain management and daily functioning). Again, all patients set personal goals on the abovementioned items. Progress on these personal goals will be monitored by the health care provider through planned telephone contact. Patients are asked to self-monitor their own health-status [37], by filling out activity and dietary diaries.

Health care providers

A total of 14 health care providers (five physiotherapists, three occupational therapists, five specialized nurses, and one dietician) are involved in the therapy sessions. All health care providers are specialized in the management of patients with musculoskeletal disorders and have experience in teaching self-management principles to groups. Moreover, all care providers took the course ‘motivational interviewing’.

The experimental intervention will be provided by one of three physiotherapists, one of three occupational therapists, one of two specialized nurses and one dietician. In the active control group, the two group sessions are provided by one of two physiotherapists and two of three specialized nurses. The telephone contact will be provided by the specialized nurses. Assignment of health care providers to the therapy programs was done on basis of availability.

Primary outcome

Daily functioning

The primary outcome of the study is the Stanford Health Assessment Questionnaire (HAQ) Disability Index during the first year after treatment [34,38]. The HAQ is an independent patient-reported outcome questionnaire containing 20 questions, covering eight domains of activities of daily living. For each item, there is a four-level response set that is scored from 0 to 3, with higher scores indicating more disability (0 = without any difficulty; 1 = with some difficulty; 2 = with much difficulty; and 3 = unable to do). Both total scores as well as each of the sub-scores range from 0 (no disability) to 3 (severe disability). An improvement of 0.26 points is considered to be clinically relevant between group change [39]. The HAQ has been found to be more responsive for measuring functioning than the WOMAC questionnaire; a widely used in hip and knee OA [40].

Secondary outcomes

Health-related quality of life (clinical efficacy)

To assess the efficacy of the interventions on health-related quality of life (HRQoL) the RAND 36-Item Health Survey 1.0 (RAND-36) will be used [41]. Scores from the eight subscales of the RAND-36 will be aggregated into two summary scores: a Physical Component Summary (PCS) and a Mental Component Summary (MCS). This instrument has been translated and validated for use in Dutch patients [42].

Patients specific complaints (clinical efficacy)

Physical functioning assessed with the patient specific complaints questionnaire (PSK). The PSK is a patient-specific questionnaire in which the patient is asked to select three activities that (s)he perceives as problematic (activities that can easily be avoided are not allowed) and scores the severity on a 10 cm visual analogue scale (VAS) [43].

Fatigue (clinical efficacy)

Fatigue is measured with the “Subjective Fatigue” subscale of the Checklist Individual Strength (CIS) [44]. The CIS is a self-administered questionnaire assessing 20 items, concerning 4 subscales divided in: subjective experienced fatigue (8 items), concentration (5 items), motivation (4 items) and physical activity (3 items). The outcomes per question are given in a 7-point scale, ranging from the statement ‘totally right’ to the statement ‘totally wrong’. The total score is counted in points with a range of 1-7 per question and a total score range of 8-56 points. The CIS is a sensitive instrument with good discriminating power and reliability [44].

Health-related quality of life (health economics)

To measure the HRQoL of patients for the purpose of economic evaluation the EuroQol-5D (EQ-5D) will be used [45]. This HRQoL instrument will be completed by the patients and is available in a validated Dutch translation. The EQ-5D is a generic HRQoL instrument comprising five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The EQ-5D index is obtained by applying predetermined weights to the five domains. This index gives a societal-based global quantification of the patient’s health status on a scale ranging from 0 (death) to 1 (perfect health). The utility weights captured by these preferences will enable the derivation of the Quality Adjusted Life Years (QALY) for each intervention and will be used in cost-utility analyses.

Patients will also be asked to rate their overall HRQoL on a visual analogue scale (EQ- 5D VAS) consisting of a vertical line ranging from 0 (worst imaginable health status) to 100 (best imaginable).

Costs (health economics)

Volumes of care will be measured prospectively using patient-based diaries (complemented by patient chart data if necessary). Per arm (intervention and control) full cost-prices will be determined using an activity based costing approach. Productivity losses for patients will be estimated using a postal questionnaire on a 3 months recall basis. The friction cost-method will be applied following the Dutch guidelines for cost analysis (Oostenbrink et al., CVZ 2004). Also travel time to a session or outpatient clinic and related costs patients make will be considered (also on the basis of 3 months recall).

The second part of the cost analysis consists of determining the cost prices for each unit of consumption in order to use these for multiplying the volumes registered for each participating patient. The Dutch guidelines for cost analyses will be used (CVZ, Oostenbrink et al., 2004). For units of care/resources where no guideline or standard prices are available real cost prices will be determined.

Study endpoints

Participants will receive postal questionnaires at baseline, and at 6, 13, 26, 39 and 52 weeks after the start of the intervention. The primary endpoint to study the long term effects of the interventions is the averaged HAQ-score [34] as obtained from the 6, 26 and 52 week time points. The 6 week time point will provide a secondary endpoint for investigating the short term effects of the interventions. On time points 6, 13, 26, 39 and 52 costs will be assessed.

Table 1. Outcome measures used at baseline and follow-up assessments.

Outcome (Instrument)	Time Points (weeks)					
	0	6	13	26	39	52
Daily functioning (HAQ) [34]	X	X	X	X	X	X
HRQoL (RAND-36) [42]	X	X		X		X
HRQoL (EuroQol EQ-5d) [45]	X	X	X	X	X	X
Patient specific complaints (PSK) [43]	X	X		X		X
Fatigue (CIS-8) [44]	X	X		X		X
Physical activity (SQUASH) [48]	X	X		X		X
Illness Cognitions (ICQ) [65]	X	X		X		X
Patient Global Assessment (PGA) [50]		X		X		X
Self-efficacy (GSES) [51]	X	X		X		X
Kinesiophobia (TSK) [66]	X	X		X		X
Costs		X	X	X	X	X

Abbreviations: HRQoL = Health related quality of life.

Socio-demographic information will be collected at baseline including age, gender, employment nature and body mass index. In Table 1 we outlined all outcome measures that will be collected at baseline and at follow-up evaluations.

Other outcomes

Since no validated outcome measures are available yet for the assessment of health status in patients with GOA, we decided to evaluate effectiveness also with a responder analysis. We developed an adapted version of the OMERACT-OARSI Responder Criteria as the secondary outcome measure of our study [46]. This composite index permits presentation of results of symptom modifying clinical trials in OA based on individual patient responses (responder yes/no). In this study, patients are considered responders if at least 3 of the 6 targeted areas (i.e. physical functioning, pain, fatigue, physical activity, acceptance, and patient global assessment) improve by $\geq 20\%$ [47]. We assess the targeted areas with the following secondary outcome measures: RAND-36-pain, PSK, CIS and SQUASH, ICQ, PGA (as described below).

Physical activity

The Short QUestionnaire to ASsess Health-enhancing physical activity (SQUASH) [48] will be used to measure physical activity. The SQUASH measures habitual physical activity level and is structured in a way that allows comparing the results to international physical activity guidelines. The questions are prestructured into activities at work, activities to/from work, household activities, leisure-time activities and sports activities. Spearman correlation has shown an overall reproducibility of 0.58 ($p < 0.05$) for the SQUASH. The SQUASH has been validated using an accelerometer, the CSA Inc. Activity Monitor (model AM7164-2.2), showing a Spearman correlation coefficient between CSA readings and total activity score of 0.45 (95% CI 0.17–0.66) [48].

Illness cognitions

Illness cognitions (acceptance and helplessness) are measured using the Illness Cognitions Questionnaire (ICQ). The ICQ is an 18-item questionnaire measuring three generic illness cognitions: helplessness, acceptance and disease benefits. Participants rate the extent to which they agree with the statements on a 4-point Likert scale, ranging from 1 (not at all) to 4 (completely). Higher scores at subscales reflect higher levels of agreement with that generic illness cognition. The scale has excellent construct and internal validity [49]. In this study we use the subscales acceptance and helplessness.

Global perceived improvement

Patient Global Assessment (PGA) is assessed by patients on a 8-point scale (1=vastly worsened; 8=completely recovered) [50].

Self-efficacy

Self-efficacy is evaluated with the General Self-Efficacy Scale (GSES) - the Dutch Language Version; a self-administered questionnaire assessing 10 items, concerning problems in daily living and the capability to bring up solutions for these problems [51]. The scores of the questions are rated on a 4-point scale. Possible responses are not at all true, hardly true, moderately true and exactly true yielding a total score between 10 and 40 points. A higher score represents a higher level of self-efficacy. The GSES was found to be configurally equivalent across 28 nations, and it forms only one global dimension. High reliability, stability, and construct validity of the GSES were confirmed in earlier studies [51].

Randomization, allocation concealment and blinding

Participants included in the study are randomly assigned to one of the treatment programmes. Restricted randomization will be used by randomly varied block sizes (2 to 6) [52]. A computer-generated randomization sequence table will be produced with random

allocation software [53] by an independent researcher (DJ). Subsequently an independent person will assign patients to one of the treatment groups. This person has no information about the persons included and has no influence on the assignment sequence or on the decision about eligibility of patients.

Patients and health care providers allocated to the experimental and active control group will be aware of the allocated arm, whereas the outcome assessor and data analysts will be kept blinded to the allocation.

Sample size

For the sample size calculation we used the statistical package G*power 3.0.10 [54]. We utilized the equation for sample size required per group using an unpaired t-test to compare differences between two independent means. To detect a minimal clinically important change of 0.26 points [39] in mean HAQ scores between both groups, assuming a SD of: 0.66 ($SE \cdot \sqrt{N} = 0.04 \cdot \sqrt{271} = 0.66$) [40] with 80% power and a two-sided 5% level, we will need 102 patients per arm (effect size is $0.26/0.66 = 0.39$). The abovementioned sample size calculation is relevant for analyses with independent t-tests.

In our analyses, however, we will use the baseline HAQ as a covariate. By a straightforward generalisation of the method described in Borm et al [55], it can be shown that in this case the sample size must be multiplied by $(1-(k-1)p)/k - \rho_B^2$ (see note below), where k is the number of follow-up assessments (3 in our case), where ρ_B is the correlation between the outcome measured at baseline and at follow-up and where p is the correlation between the follow-up measurements. Although some publications report test-rest correlation for the HAQ of more than 0.8 (31), there is no direct information about the correlation that is to be expected in our trial. We expect ρ_B to be smaller than p (even within the treatment groups), because the interventions will be ‘between’ the baseline and follow-up assessments. The interventions may not have the same effect on all patients and therefore decrease the correlation. When p is between 0.7 and 0.9 and $\rho_B = p - 0.2$, the sample size can be reduced by a (design) factor 0.44 to 0.55. For p between 0.8 and 0.9 and $\rho_B = p - 0.1$, the sample size can be reduced by a factor 0.38 to 0.44. A trial with 55 patients per treatment group will then have at least 80% power (when the design factor is 0.55). In the most optimistic scenario, when the design factor is 0.38, the study has slightly over 90% power.

Finally, as the patients will be treated in groups (cluster) of approximately eight, the patient numbers have to be increased by a factor $1+(8-1) \cdot ICC$. For $ICC=0.05$, this leads to 74 patients. In order to compensate for possible drop-outs (15%), we plan to enrol 85 patients per treatment group.

Note: When $Y = \text{mean}(Y_1, \dots, Y_k)$ and the baseline and follow-up measurements Y_i all have standard deviation σ , the variance of Y is $((1+(k-1)p)/k)\sigma^2$, the correlation between B and Y

is $k \rho_B / \sqrt{k + (k-1) \rho}$ and the formula in Borm *et al* yields the design factor $(1 - (k-1)\rho)/k - \rho_B^2$.

Planned data analysis

Study data are entered in Access 2003, exported to the statistical package STATA v10 stored on a secure network drive. Five percent of the data will be entered twice to assess percentage and nature of typing errors. All paper records are stored in a locked cabinet in an anonymised format. The researcher will check for any missing data and will manage this according to the recommendations of the questionnaires. Descriptive statistics will be used to determine participant characteristics. Continuous variables will be reported using means, standard deviations (SD) and inter-quartile ranges when appropriate, if not median and ranges are shown. For dichotomous/categorical variables, we will display absolute numbers and percentages. The primary analysis will be according to the intention to treat principle.

Clinical efficacy

The primary variable, HAQ during the first year after treatment, will be analysed with a random effects model with the HAQ scores after 6, 26 and 52 weeks as dependent variable. The fixed factors will be assessment (6, 26 or 52 weeks), treatment group, sex and baseline value. In order to account for the group wise treatment and the repeated measurements, the random effects group and patient will be included. HAQ immediately after treatment will be evaluated in a random effects model with fixed factors treatment group, sex and baseline value, and random factor group. All other continuous variables will be analysed in a similar way. Skewed variables will be transformed before analysis. For dichotomous outcomes, random effects general linear models with Bernoulli distribution and linear link function will be used, similar to the ones described for continuous outcomes. Changes in effect size over time will be evaluated by adding the interaction of assessment and treatment group to the model.

Health economics

The economic evaluation is based on the general principles of a cost-effectiveness analysis and cost-utility analysis. For the cost-effectiveness analysis we will calculate the incremental cost-effectiveness ratio (ICER) as cost per unit improvement on the HAQ. For the cost-utility analysis we will calculate the ICER as cost per Qaly gained. This ICER will be evaluated stochastically and uncertainty will be determined using the bootstrap method and/or Fieller method. A cost-effectiveness acceptability curve will be derived that is able to evaluate efficiency by using different thresholds (Willingness To Pay) for a QALY. The impact of uncertainty surrounding deterministic parameters (for example cost-prices) on the ICER will be explored using one-way sensitivity analyses on the range of extremes. The

economic evaluation is done along-side the clinical trial and consequently adheres to the earlier presented design.

DISCUSSION

To date, research on NPT options for OA has mainly focused upon patients with hip and knee OA. In 2008, NICE disseminated multiple recommendations for future OA research based on the research hiatuses they identified. One of their research questions was: “What are the benefits of individual and combination OA therapies in people with multiple joint region pain?”. This study will contribute to the body of evidence on NPT in GOA patients.

A possible limitation in our study is the limited contrast in the content of the experimental and control intervention. Both interventions were developed from a clinical and pragmatic perspective. Since both interventions should be directly implementable in clinical practice after study completion, we created two treatment protocols according to the recommendations outlined in OA guidelines and current best-practice. The content of both interventions is very similar but several critical differences are apparent such as the mode of delivery, the number of involved health care providers and the number of group-sessions. Specific insights in the effectiveness and costs of these differences will aid health care providers and care vendors in their decision making for the management of patients with GOA.

To our knowledge we are the first to define GOA from a clinical rather than a radiographic perspective, as no consistent clinical useful definition of GOA is available. In 1952, Kellgren and Moore defined GOA as involvement of multiple joints combined with Heberden’s nodule [31]. Since then, multiple definitions of GOA have been used, for the greater part based on radiological changes. Most definitions state that GOA involves at least three joints [31], although this again is questioned [27]. The group of joints most often incorporated in definitions are the hands, neck, lower back, knees and hips [7,56,57], whereas other definitions postulate that the involvement of atypical joints [25,26] or hallux valgus [25,58] is essential for GOA. To date, two specific phenotypes of GOA have been established [28], however, these phenotypes are far from useful in daily practice as these phenotypes only represent a very small proportion of patients with OA-like complaints in multiple joints. Considering the low feasibility and desirability of obtaining radiographs of a large number of joint in clinical practice, we believe that clinical signs and symptoms should also be taken into account in the definition of GOA. Especially, since pain at multiple joint sites is associated with lower levels of functioning [21,23,59-61], more pain [21,23,59,60] and higher levels of distress [21,59,62], and complaints rather than radiographic OA are the main motivation for patients to engage in therapy. So, for the purpose of this project we formulated a pragmatic definition of GOA (as described earlier)

based upon literature findings and on the basis of consensus of several clinicians and health professionals with experience with patients with GOA.

There is a need for outcome measures to evaluate self-management interventions [63]. Self-management is defined as the individual's ability to manage the symptoms, treatment, physical and psychosocial consequences and lifestyle changes inherent in living with a chronic condition [64]. Characteristically, one or more of these areas are addressed by self-management interventions [63]. In our study we target physical functioning, pain, fatigue, physical activity, and acceptance. However, no comprehensive outcome measure is available to measure all these different aspects. Mulligan *et al* (2005) state that when designing a self-management intervention, it is important to be clear about what the intervention is designed to achieve, in what areas it is likely to have an effect, and to choose outcome measures accordingly [63]. Therefore, we decided to include a responder analysis - derived from the OMERACT-OARSI responder criteria [46] – as one of the secondary measures in our analysis that specifically evaluates those areas we aim to address. In a future publication we intend to evaluate and discuss this method of assessing self-management interventions.

In conclusion, this study will provide additional insights in the effectiveness of non-pharmacological interventions for GOA. The publication of our study protocol enables future readers to compare what was originally intended with what was actually done, thus preventing both “data dredging” and post-hoc revisions of study aims.

REFERENCE LIST

- (1) Petersson IF, Jacobsson LT: Osteoarthritis of the peripheral joints. *Best Pract Res Clin Rheumatol* 2002, 16: 741-760.
- (2) Cooper C, Snow S, McAlindon TE, Kellingray S, Stuart B, Coggon D *et al.*: Risk factors for the incidence and progression of radiographic knee osteoarthritis. *Arthritis Rheum* 2000, 43: 995-1000.
- (3) Felson DT, Zhang Y: An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. *Arthritis Rheum* 1998, 41: 1343-1355.
- (4) Felson DT: An update on the pathogenesis and epidemiology of osteoarthritis. *Radiol Clin North Am* 2004, 42: 1-9, v.
- (5) Hart DJ, Doyle DV, Spector TD: Incidence and risk factors for radiographic knee osteoarthritis in middle-aged women: the Chingford Study. *Arthritis Rheum* 1999, 42: 17-24.
- (6) Garstang SV, Stitik TP: Osteoarthritis: epidemiology, risk factors, and pathophysiology. *Am J Phys Med Rehabil* 2006, 85: S2-11.
- (7) Gunther KP, Sturmer T, Sauerland S, Zeissig I, Sun Y, Kessler S *et al.*: Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. *Ann Rheum Dis* 1998, 57: 717-723.
- (8) Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N *et al.*: OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage* 2008, 16: 137-162.
- (9) Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N *et al.*: OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthritis Cartilage* 2007, 15: 981-1000.
- (10) Chodosh J, Morton SC, Mojica W, Maglione M, Suttrop MJ, Hilton L *et al.*: Meta-analysis: chronic disease self-management programs for older adults. *Ann Intern Med* 2005, 143: 427-438.
- (11) Warsi A, LaValley MP, Wang PS, Avorn J, Solomon DH: Arthritis self-management education programs: a meta-analysis of the effect on pain and disability. *Arthritis Rheum* 2003, 48: 2207-2213.
- (12) Deyle GD, Henderson NE, Matekel RL, Ryder MG, Garber MB, Allison SC: Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. A randomized, controlled trial. *Ann Intern Med* 2000, 132: 173-181.
- (13) Deyle GD, Allison SC, Matekel RL, Ryder MG, Stang JM, Gohdes DD *et al.*: Physical therapy treatment effectiveness for osteoarthritis of the knee: a randomized comparison of supervised clinical exercise and manual therapy procedures versus a home exercise program. *Phys Ther* 2005, 85: 1301-1317.
- (14) Fransen M, Crosbie J, Edmonds J: Physical therapy is effective for patients with osteoarthritis of the knee: a randomized controlled clinical trial. *J Rheumatol* 2001, 28: 156-164.

- (15) Roddy E, Zhang W, Doherty M: Aerobic walking or strengthening exercise for osteoarthritis of the knee? A systematic review. *Ann Rheum Dis* 2005, 64: 544-548.
- (16) Weinberger M, Tierney WM, Cowper PA, Katz BP, Booher PA: Cost-effectiveness of increased telephone contact for patients with osteoarthritis. A randomized, controlled trial. *Arthritis Rheum* 1993, 36: 243-246.
- (17) Rene J, Weinberger M, Mazzuca SA, Brandt KD, Katz BP: Reduction of joint pain in patients with knee osteoarthritis who have received monthly telephone calls from lay personnel and whose medical treatment regimens have remained stable. *Arthritis Rheum* 1992, 35: 511-515.
- (18) Weinberger M, Tierney WM, Booher P, Katz BP: Can the provision of information to patients with osteoarthritis improve functional status? A randomized, controlled trial. *Arthritis Rheum* 1989, 32: 1577-1583.
- (19) Eakin E, Reeves M, Lawler S, Graves N, Oldenburg B, Del MC *et al.*: Telephone counseling for physical activity and diet in primary care patients. *Am J Prev Med* 2009, 36: 142-149.
- (20) Bulthuis Y, Drossaers-Bakker KW, Taal E, Rasker J, Oostveen J, van't Pad BP *et al.*: Arthritis patients show long-term benefits from 3 weeks intensive exercise training directly following hospital discharge. *Rheumatology (Oxford)* 2007, 46: 1712-1717.
- (21) Croft P, Jordan K, Jinks C: "Pain elsewhere" and the impact of knee pain in older people. *Arthritis Rheum* 2005, 52: 2350-2354.
- (22) Kamaleri Y, Natvig B, Ihlebaek CM, Bruusgaard D: Localized or widespread musculoskeletal pain: does it matter? *Pain* 2008, 138: 41-46.
- (23) Peat G, Thomas E, Wilkie R, Croft P: Multiple joint pain and lower extremity disability in middle and old age. *Disabil Rehabil* 2006, 28: 1543-1549.
- (24) Splawinski J, Kuzniar J: Clinical trials: active control vs placebo--what is ethical? *Sci Eng Ethics* 2004, 10: 73-79.
- (25) Carroll GJ: Polyarticular osteoarthritis--two major phenotypes hypothesized. *Med Hypotheses* 2006, 66: 315-318.
- (26) Ala-Kokko L, Baldwin CT, Moskowitz RW, Prockop DJ: Single base mutation in the type II procollagen gene (COL2A1) as a cause of primary osteoarthritis associated with a mild chondrodysplasia. *Proc Natl Acad Sci U S A* 1990, 87: 6565-6568.
- (27) Cooper C, Egger P, Coggon D, Hart DJ, Masud T, Cicuttini F *et al.*: Generalized osteoarthritis in women: pattern of joint involvement and approaches to definition for epidemiological studies. *J Rheumatol* 1996, 23: 1938-1942.
- (28) Carroll GJ, Bredahl WH, Jazayeri J: Confirmation of two major polyarticular osteoarthritis (POA) phenotypes--differentiation on the basis of joint topography. *Osteoarthritis Cartilage* 2009, 17: 891-895.
- (29) Dougados M, Nakache JP, Gueguen A: Criteria for generalized and focal osteoarthritis. *Rev Rhum Engl Ed* 1996, 63: 569-575.
- (30) Loughlin J, Irlen C, Fergusson C, Sykes B: Sibling pair analysis shows no linkage of generalized osteoarthritis to the loci encoding type II collagen, cartilage link protein or cartilage matrix protein. *Br J Rheumatol* 1994, 33: 1103-1106.
- (31) Kellgren JH, Moore R: Generalized osteoarthritis and Heberden's nodes. *Br Med J* 1952, 1: 181-187.

- (32) Riyazi N, Meulenbelt I, Kroon HM, Runday KH, Hellio le Graverand MP, Rosendaal FR *et al.*: Evidence for familial aggregation of hand, hip, and spine but not knee osteoarthritis in siblings with multiple joint involvement: the GARP study. *Ann Rheum Dis* 2005, 64: 438-443.
- (33) Vignon E: Hand osteoarthritis and generalized osteoarthritis: a need for clarification. *Osteoarthritis Cartilage* 2000, 8 Suppl A: S22-S24.
- (34) Fries JF, Spitz P, Kraines RG, Holman HR: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980, 23: 137-145.
- (35) Perera R, Heneghan C, Yudkin P: Graphical method for depicting randomised trials of complex interventions. *BMJ* 2007, 334: 127-129.
- (36) Glasgow RE, Emont S, Miller DC: Assessing delivery of the five 'As' for patient-centered counseling. *Health Promot Int* 2006, 21: 245-255.
- (37) Glasziou P, Irwig L, Mant D: Monitoring in chronic disease: a rational approach. *BMJ* 2005, 330: 644-648.
- (38) Boers M, Jacobs JW, van V, V, van Riel PL: Consensus Dutch health assessment questionnaire. *Ann Rheum Dis* 2007, 66: 132-133.
- (39) Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE, Jr.: Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum* 2000, 43: 1478-1487.
- (40) Bruce B, Fries J: Longitudinal comparison of the Health Assessment Questionnaire (HAQ) and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). *Arthritis Rheum* 2004, 51: 730-737.
- (41) Blanchard CM, Cote I, Feeny D: Comparing short form and RAND physical and mental health summary scores: results from total hip arthroplasty and high-risk primary-care patients. *Int J Technol Assess Health Care* 2004, 20: 230-235.
- (42) VanderZee KI, Sanderman R, Heyink JW, de HH: Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. *Int J Behav Med* 1996, 3: 104-122.
- (43) Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W *et al.*: A patient-specific approach for measuring functional status in low back pain. *J Manipulative Physiol Ther* 1999, 22: 144-148.
- (44) Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G: Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994, 38: 383-392.
- (45) The EuroQol Group: EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990, 16: 199-208.
- (46) Pham T, van der HD, Lassere M, Altman RD, Anderson JJ, Bellamy N *et al.*: Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. *J Rheumatol* 2003, 30: 1648-1654.
- (47) Philadelphia Panel: Evidence-based clinical practice guidelines on selected rehabilitation interventions: overview and methodology. *Phys Ther* 2001, 81: 1629-1640.
- (48) Wendel-Vos GC, Schuit AJ, Saris WH, Kromhout D: Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. *J Clin Epidemiol* 2003, 56: 1163-1169.

- (49) Maas M, Taal E, van der LS, Boonen A: A review of instruments to assess illness representations in patients with rheumatic diseases. *Ann Rheum Dis* 2009, 68: 305-309.
- (50) Bellamy N: Science of assessment. *Ann Rheum Dis* 2005, 64 Suppl 2: ii42-ii45.
- (51) Luszczyńska A, Scholz U, Schwarzer R: The general self-efficacy scale: multicultural validation studies. *J Psychol* 2005, 139: 439-457.
- (52) Altman DG, Bland JM: How to randomise. *BMJ* 1999, 319: 703-704.
- (53) Saghaei M: Random allocation software for parallel group randomized trials. *BMC Med Res Methodol* 2004, 4: 26.
- (54) Faul F, Erdfelder E, Lang AG, Buchner A: G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007, 39: 175-191.
- (55) Borm GF, Fransen J, Lemmens WA: A simple sample size formula for analysis of covariance in randomized clinical trials. *J Clin Epidemiol* 2007, 60: 1234-1238.
- (56) Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM: Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis Rheum* 1995, 38: 1134-1141.
- (57) Cerhan JR, Wallace RB, el-Khoury GY, Moore TE: Risk factors for progression to new sites of radiographically defined osteoarthritis in women. *J Rheumatol* 1996, 23: 1565-1578.
- (58) Roddy E, Zhang W, Doherty M: Prevalence and associations of hallux valgus in a primary care population. *Arthritis Rheum* 2008, 59: 857-862.
- (59) Carnes D, Parsons S, Ashby D, Breen A, Foster NE, Pincus T *et al.*: Chronic musculoskeletal pain rarely presents in a single body site: results from a UK population study. *Rheumatology (Oxford)* 2007, 46: 1168-1170.
- (60) Dawson J, Linsell L, Zondervan K, Rose P, Randall T, Carr A *et al.*: Epidemiology of hip and knee pain and its impact on overall health status in older adults. *Rheumatology (Oxford)* 2004, 43: 497-504.
- (61) Keenan AM, Tennant A, Fear J, Emery P, Conaghan PG: Impact of multiple joint problems on daily living tasks in people in the community over age fifty-five. *Arthritis Rheum* 2006, 55: 757-764.
- (62) Kamaleri Y, Natvig B, Ihlebaek CM, Benth JS, Bruusgaard D: Number of pain sites is associated with demographic, lifestyle, and health-related factors in the general population. *Eur J Pain* 2008, 12: 742-748.
- (63) Mulligan K, Newman SP, Taal E, Hazes M, Rasker JJ: The design and evaluation of psychoeducational/self-management interventions. *J Rheumatol* 2005, 32: 2470-2474.
- (64) Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J: Self-management approaches for people with chronic conditions: a review. *Patient Educ Couns* 2002, 48: 177-187.
- (65) Evers AW, Kraaijaat FW, Geenen R, Bijlsma JW: Psychosocial predictors of functional change in recently diagnosed rheumatoid arthritis patients. *Behav Res Ther* 1998, 36: 179-193.
- (66) Vlaeyen JW, Kole-Snijders AM, Boeren RG, van EH: Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 1995, 62: 363-372.

9

Piloting and preliminary evaluating a multi-disciplinary care program for individuals with generalized osteoarthritis on pain and self-efficacy: a concurrent multiple-baseline single case study.

ABSTRACT

Objectives. To evaluate the feasibility and potential effectiveness of a 12-week non-pharmacological multidisciplinary intervention in patients with generalised osteoarthritis (GOA).

Design. A randomised, concurrent, multiple-baseline, single-case design. During the baseline period, the intervention period, and the post-intervention period, all participants completed several health outcomes twice a week on visual analogue scales.

Setting. Rheumatology, outpatient department of a specialized hospital in the Netherlands.

Participants. One man and four women (age 51 to 76) diagnosed with GOA.

Primary outcome measures. To assess feasibility we assessed the number of drop-outs and adverse events, adherence rates, and patient satisfaction.

Secondary outcome measures. To assess the potential effectiveness we assessed pain and self-efficacy using visual data inspection and randomisation tests.

Results. The intervention was feasible in terms of adverse events (none) and adherence rate, but not in terms of participant satisfaction with the intervention. Visual inspection of the data and randomisation testing demonstrated no effects on pain ($p = 0.93$) or self-efficacy ($p = 0.85$).

Conclusions. The results of the present study indicate that the proposed intervention for patients with GOA was insufficiently feasible and effective. The data obtained through this multiple-baseline study has highlighted several areas in which the therapy programme can be optimised.

Hoogeboom TJ, Kwakkenbos L, Rietveld L, den Broeder AA, de Bie RA, van den Ende CHM.
*Piloting and preliminary evaluating a multi-disciplinary care program for individuals with
generalized osteoarthritis on pain and self-efficacy:
a concurrent multiple-baseline single case study.*
BMJ Open. 2012 Jul 19;2(4).

INTRODUCTION

A growing body of evidence shows that individuals with established osteoarthritis with multiple joint involvement - often referred to as generalised osteoarthritis (GOA) - represent a relatively large subgroup of patients [1-4]. It has been suggested that these people might be in need of more intensive treatment options than patients with single joint complaints [1,5]. To the best of our knowledge, however, there are no studies that evaluate non-pharmacological, multidisciplinary care in individuals with GOA [5], warranting the development and evaluation of such a treatment programme. Therefore, we conceptualised a non-pharmacological, multidisciplinary treatment programme following a previously-described systematic procedure [6]. The intervention was based on recommendations for the management of hip and knee osteoarthritis [7-9], and was tailored to the needs of patients with multiple joint involvement [1]. Due to the complex nature of multiple joint-involvement in OA [1-4] and the fact that guidelines for hip and knee OA recommend multiple non-pharmacological treatment modalities, an intervention was developed by a multidisciplinary team [8].

Before evaluating such an intervention in a randomised clinical trial, a pilot study is recommended [10], since evaluations are often undermined by problems of acceptability, compliance, delivery of the intervention, recruitment and retention, and smaller-than-expected effect sizes [11]. A useful study design for pilot interventions is the multiple-baseline single-case design, as it allows researchers to test the feasibility of the intervention and to make an assessment of its potential effectiveness with a low number of participants [12]. In a multiple-baseline design, the intervention is introduced to subjects after randomly-assigned baseline periods of different lengths, and an effect is demonstrated if the measured outcome only changes after the intervention has been introduced [13].

The primary aim of our study was to evaluate the feasibility of a non-pharmacological, multidisciplinary intervention in patients with GOA. Our secondary aim was to assess the potential effectiveness of this intervention on pain and self-efficacy.

METHODS

Participants

Men and women, 40 years or older and referred to the multidisciplinary intervention, were eligible to participate in the present study if they had been diagnosed with GOA; i.e. experiencing complaints in three or more joint groups, having at least two objective signs that indicate OA in at least two joints, and having limitations in daily functioning (Health Assessment Questionnaire-Disability Index score (HAQ-DI) [14] > 0.5) [15]. Individuals

were excluded from participation in the intervention if: 1) they were awaiting joint replacement surgery, 2) they had already participated unsuccessfully in a self-management programme for their GOA complaints, 3) their therapists suspected that they were suffering high levels of distress, 4) they did not master the Dutch language, or 5) they were illiterate. Recruitment and treatment of patients took place at the rheumatology outpatient department at the Maartenskliniek Woerden (the Netherlands). The study protocol was reviewed and approved by the Institutional IRB of the UMC Nijmegen (protocol number 2009/173), and did not fall within the remit of the Medical Research Involving Human Subjects Act.

Design

A randomised concurrent multiple-baseline single-case design was applied [13]. Participants completed repeated measurements during a baseline phase (phase A), a therapy-phase (phase B, 12 weeks) and a post-therapy phase (phase A'). Phase A acted as a control and was therefore compared with phases B and A'. By applying multiple baselines of varying length, observed effects of the treatment can be distinguished from effects due to chance [12,16,17], thus increasing internal validity. The total duration of phase A and A' was set at 7 weeks for each participant, and consequently participants with a longer phase A had a shorter phase A'. Participants were randomly assigned to a baseline and post-therapy period of either 2 and 5 weeks, 2.5 and 4.5 weeks, 3 and 4 weeks, ..., or 5 and 2 weeks, respectively, using the Wampold-Worsham method [18] to increase statistical power. During the total study period of 19 weeks, participants completed diary measures twice a week, resulting in a total of 38 measurement points (14 during phase A and A' and 24 during phase B). Each diary measure comprised 14 VAS scales.

Measurements

Feasibility of the intervention

To evaluate the feasibility of the intervention, we assessed: 1) number of, and reasons for, drop-out during the intervention; 2) adherence to the intervention (number of no shows); 3) occurrence of adverse events related to the intervention; 4) participants' satisfaction with the intervention (straightforward question ranging from 0 (totally dissatisfied) to 10 (totally satisfied)); and 5) participants' satisfaction with the assessment procedure (straightforward yes/no questions).

Diary Measures

Diary measures comprised 14 VAS scales (scoring range from 0 to 10). Pain and fatigue were measured by single straightforward questions. Furthermore, 12 items derived from validated questionnaires were scored on a VAS scale. Kinesiophobia was measured with four VAS scales [19]. Self-efficacy was assessed using two questions from the Arthritis Self-Efficacy Scale [20]. Acceptance of the disease was measured with two questions from

the subscale Acceptance of the Illness Cognition questionnaire [21], and illness perceptions were evaluated by two questions from the Illness Perception questionnaire [22]. To assess the specific complaints of each participant, we used the Patient-Specific Complaints questionnaire (PSK) [23]. The most important complaint was assessed through the diary measure. For all scales, a higher score represented unfavourable outcomes. Pain and self-efficacy were our main secondary outcome measures.

Pre- and post-intervention measures

At baseline, we collected data on age, sex, level of education (low (no or primary education), medium (secondary school and/or preparatory middle-level vocational education), high (university of applied sciences and/or university)) and duration of symptoms. Prior to the start of the programme, we also assessed participant's expectations about its effectiveness on a scale from 0 to 10 (0 representing 'No expectations whatsoever'). Pre- and post-intervention measures consisted of a set of validated questionnaires. We measured fatigue with the "Subjective Fatigue" subscale of the Checklist Individual Strength (CIS) [24], on which higher scores represent greater fatigue. Self-efficacy was evaluated with the General Self-Efficacy Scale [25], where higher scores represent higher levels of self-efficacy. Acceptance and helplessness were measured using the Illness Cognitions Questionnaire (ICQ) [26], where higher scores reflect higher levels of agreement with that generic illness cognition. As no specific questionnaires are available to assess the self-reported functional status of individuals with GOA, we used generic questionnaires for both the lower and upper extremities, namely the Lower Extremity Functional Scale (LEFS) [27] and the Disability of Arm, Shoulder and Hand (DASH), respectively [28]. Higher scores on the LEFS and DASH represent lower and greater disability, respectively.

Intervention

The group-based intervention (8 persons per group) lasted 12 weeks, comprised 10 sessions of approximately 1.5 hours per session, and was provided by an occupational therapist and physical therapist. To ensure group learning the treatment program was decided to be delivered in a group setting. The intervention aimed to increase the participants' knowledge of the disease, to optimise the participants' current lifestyle, and to enhance the participants' self-efficacy in controlling the disease. To do so, patients received information on activity pacing, medication use, physical activity and weight reduction. Consequently, based on the received information, participants set personal goals regarding all these health areas. By setting these personal goals, participants transferred the health information into practical and personally relevant therapy goals. Goal setting and monitoring was done according to the 5-As model of behaviour change counselling [29]; a generally accepted method to enhance self-efficacy in health care settings. During each session, after the initial

information session, the individual goals were monitored and discussed. To allow for positive feedback regarding the personal goals, all goals had to be achievable in brief amounts of time. Some examples of personal therapy goals were: 1. For the next three days, while at work, plan and perform 15 minutes of physical activity spread over three different time points (component Physical Activity); 2. For the next week, whilst cleaning the house, alternate (maximum of 10 minutes) between vacuum cleaning, other household chores, and rest moments (component Activity Pacing); 3. For the next week, use your pain medication (two tablets of Paracetamol (500 mg)) four times a day and monitor your pain during this period (component Medication Use); and 4. For the next week, eat at least three days two slices of whole wheat bread as breakfast (component Weight Reduction). In addition, daily activities (such as walking, sitting, standing, stair climbing and getting in and out of bed) were included in the therapeutic activity programme. Participants received information and practised how to perform these daily activities without overexerting the joints and muscles. Participants were instructed and encouraged to implement these techniques and methods of performing the activities in their daily practice. Finally, participants were familiarised with different kinds of sports, tailored to the participants' complaints to prevent overexertion (i.e. tai chi, brisk walking, and therapeutic fitness). An overview of the intervention is depicted in Box 1. Participants were advised to implement these recommendations in their home situation.

Data analysis

All data were entered into the data-entry program Epidata [30]. Ten per cent of the data was entered twice to establish the quality of data entry. Missing data were described. Diary data were analysed using the 2-Standardised Deviation (SD) band method [17] (visual inspection) and randomisation tests [31]. The 2-SD band was calculated from the baseline data and graphed from the baseline phase through the intervention phase. If two or more successive data points in the intervention or post-intervention phase fell outside the bandwidth of 2 SDs, the result was considered significant [17]. As serial dependence - the extent to which scores at one point in a series are predictive of scores at another point in the same data set - can bias the visual inspection [17], we checked our data in each phase for serial dependence using the lag-1 method [12]. If data were found to be significantly correlated, we transformed the data using a moving-average transformation, in which the preceding and succeeding measurements were taken into account [12,16]. In addition, randomisation tests for multiple-baseline single-case designs were carried out. We expected phase B and A' to be superior to phase A in terms of our health outcome assessment. Therefore we tested the null hypothesis - that there would be no differential effect for any of the measurement times - using a randomisation test of the differences in the means between the pre-intervention phase and the intervention or post-intervention phase [17]. A p -value < 0.05 was considered statistically significant.

Box 1. Pat-plot of the multidisciplinary intervention.

Timeline		Intervention
Pre-measurement		
Week 1	Part 1	a
	Part 2	b c
Week 2	Part 1	d e
	Part 2	b c
Week 3	Part 1	d ①↻
	Part 2	c ①↻ ②↻
Week 4		e ①↻ ②↻
Week 5		b d ①↻
Week 6		c ①↻ ②↻
Week 7		e ①↻
Week 9		b c ①↻
Week 12		d ①↻ f
Post-measurement		

a	Introduction meeting.
b	Information on Pain and Medication use.
c	Information on Activity Pacing.
d	Information on the importance of Physical Activity.
e	Information on Weight Reduction.
①↻	Activity programme to improve quality of movement.
②↻	Sports activity.
f	Evaluation time point.

For the pre- and post-measurements, we considered change scores of 20% on validated questionnaires as clinically relevant [32]. We used Stata/IC 10.1 for Windows for the descriptive and visual analysis of the data and R version 2.14.1 for the randomisation tests [31].

RESULTS

Nine people were screened to participate in the study; two patients were excluded as they did not report functional disabilities (HAQ-DI < 0.5) and two patients who were eligible were unable to attend the program. Eventually, five participants gave written informed consent to participate in the study. One patient dropped out of the study within two weeks after the start of the study, reporting that filling out the questionnaires was too demanding for her on an emotional level. However, she did continue with the multidisciplinary intervention. The four remaining participants completed all 38 diary measures, resulting in 2,128 completed items. Six items (0.3%) were missing. Data entry errors were negligible (<0.1%). Table 1 presents the characteristics of the participants.

Table 1. Characteristics of the study participants.

Participant	Sex	Age (y)	Education	No. painful joint groups (0 - 11)	Baseline assignment (measurements)
1	F	76	Low	8	4
2	F	68	Medium	3	5
3	M	59	Low	11	7
4	F	56	High	5	6
5 [†]	F	51	High	-	6

Abbreviations: F, female; M, male; No., number of.

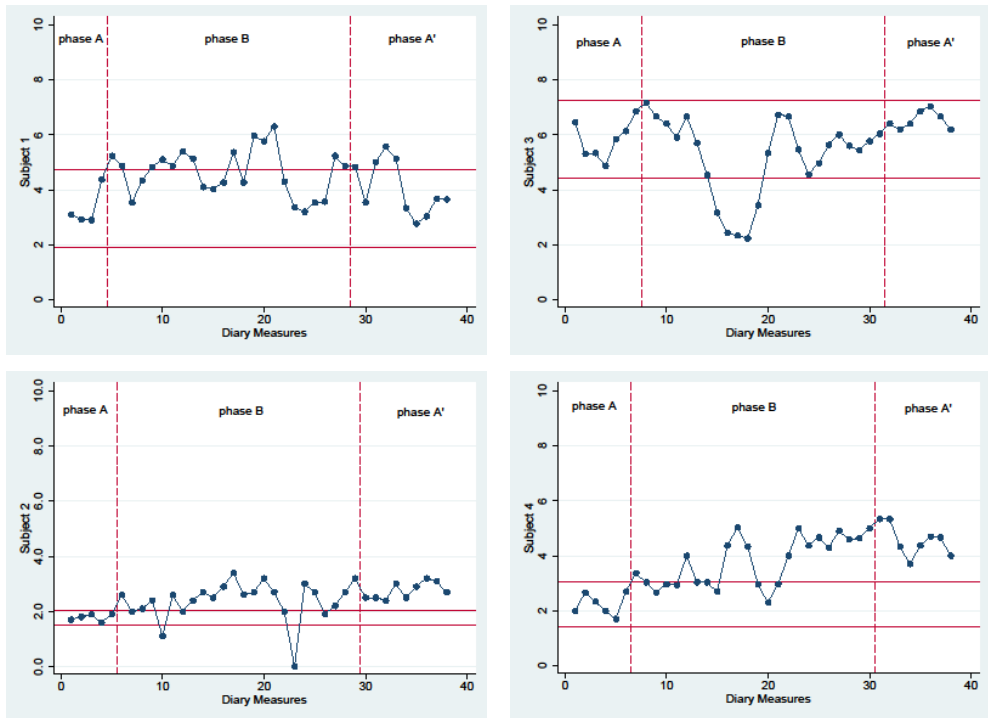
[†] Dropped out.

Feasibility of the intervention

Prior to the intervention, participants' expectations regarding the effectiveness of the intervention ranged from 5 to 7 (median = 7). Participant 3 missed three of the 10 sessions; participants 2 and 4 both missed one session. Participant 1 reported an increase in pain levels, which she ascribed to the intervention. Satisfaction with the intervention was assigned a score of 8 points out of 10 by participants 1, 2 and 4, and 7 points out of 10 by participant 3. Perceived therapy effects were assigned a score of 7, 3, 5, and 7 out of 10 by participants 1, 2, 3 and 4, respectively. All participants believed the questionnaires used in this study properly evaluated their most important issues. The remarks most frequently made by participants regarding the intervention were: 1) there were too many sessions and these were too short/brief; 2) too much verbal information; 3) too much time between two

sessions; 4) too little information on acceptance of the disease; and 5) too little individualisation in the exercise sessions, and in setting and monitoring therapy goals.

Figure 1. Diary measures for pain with 2-SD horizontal band graph for baseline (phase A), intervention (phase B) and post-intervention (phase A') phases. Scores on the pain VAS range from 0 to 10; higher scores indicate higher levels of pain.

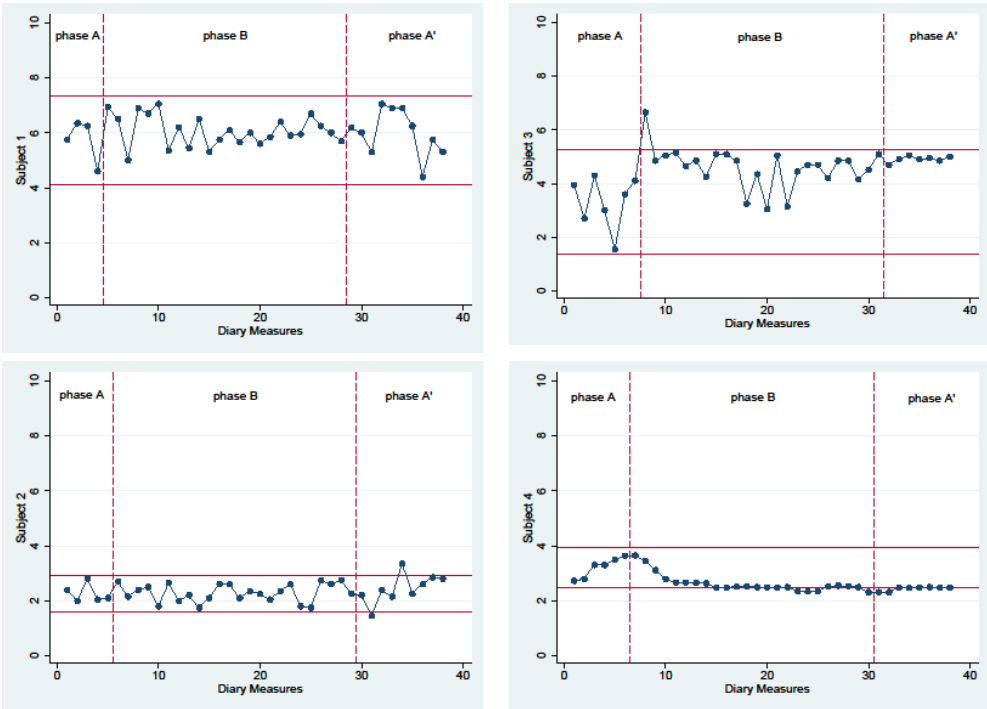


Diary measures

Our primary effectiveness outcome measures were pain and self-efficacy. In the pain data, participant 3's intervention phase showed serial dependence, and that of participants 1 and 4 showed large fluctuations. Thus, we transformed these data prior to completion of visual data analysis. The 2-SD band method showed that participants 1, 2 and 4 each experienced significant deterioration in their pain scores between baseline, intervention and post-intervention phases. Participant 3 demonstrated significant improvement during the intervention phase (Figure 1), though this did not persist during the post-intervention phase. For all four participants, randomisation tests demonstrated no significant changes in pain between the pre-intervention phase and the intervention/post-intervention phase ($p=0.93$). Serial dependence was found in participant 4's self-efficacy data, and these data were transformed prior to the analyses. The 2-SD band method demonstrated that participant 4

experienced significantly higher levels of self-efficacy in both the intervention and post-intervention phase compared to the baseline phase. No differences were found for participants 1, 2 and 3 (Figure 2).

Figure 2. Diary measures for Self-Efficacy with 2-SD horizontal band graph for baseline (phase A), intervention (phase B) and post-intervention (phase A') phases. Scores on the pain VAS range from 0 to 10, higher scores indicating lower levels of self-efficacy.



Randomisation testing demonstrated no statistically significant difference between the phase prior to the intervention and the phases during and after the intervention ($p=0.85$). Outcomes of the randomisation tests for our secondary effectiveness outcome measures were: fatigue ($p=0.79$), patient specific complaints ($p=0.64$), kinesiphobia ($p=0.02$), illness cognitions ($p=0.69$) and illness perception ($p=0.60$).

Pre- and post-measurements

Table 2 depicts the clinically relevant changes from baseline for each of the four participants. None of the participants reported improvement in self-efficacy. Participant 1 experienced clinically relevant deterioration in self-efficacy, upper body function and

kinesiophobia. Participant 4 reported improvements in fatigue levels, upper body function, kinesiophobia and acceptance. Both participants 2 and 3 remained stable.

Table 2. Clinically relevant differences between baseline and post-intervention measurements.

	Fatigue		Self-efficacy		Function				Kinesiophobia		Illness Cognitions			
					Upper		Lower				Help		Accept	
	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1	T0	T1
pt1	42	39	<u>35</u>	<u>27</u>	<u>35</u>	<u>50</u>	44	47	43	50	11	11	12	12
pt2	9	9	35	37	18	13	69	68	28	31	8	9	23	24
pt3	56	33	35	30	<u>31</u>	<u>43</u>	38	41	57	53	13	14	15	19
pt4	34	27	29	31	44	32	46	48	48	34	9	9	11	14

Bold = 20% improvement, Underlined = 20% deterioration. Abbreviations: Accept = Subscale Acceptance; Help = Subscale Helplessness; Lower = Lower extremity functioning; pt# = Participant #; T0 = Baseline measurement; T1 = Post-intervention measurement; Upper = Upper extremity functioning.

DISCUSSION

Our data suggest that the tailored, 12-week non-pharmacological, multidisciplinary intervention for patients with GOA was feasible in terms of adverse events, number of drop-outs and participation rate. On the other hand, the participants raised several critical points concerning the structure, content, and perceived benefits of the intervention. The latter was confirmed by visual inspection of the data and randomisation testing, as the intervention did not demonstrate clear-cut effects on health-related factors. Therefore, we believe the content and structure of the current intervention does not warrant further evaluation in a randomised clinical trial.

In view of the participants' remarks, we believe that the intervention should be more individually tailored. One of the remarks was that the therapeutic movement programme was not sufficiently individualised to address the participants' health problems. In a future non-pharmacological, multidisciplinary intervention, it might be of value to incorporate the results of the Patient-Specific Complaints instrument [23] in the therapeutic activity programme. Moreover, it was suggested that setting and achieving goals should be monitored more closely. To do so, participants should draw up action plans by completing goal-setting forms to formulate short-term goals, whilst being aware of potential limiting factors. In this way, personal goals could be monitored, discussed and adjusted, which in turn might increase the involvement and self-efficacy of the participants [17]. Finally, participants had relatively low treatment expectations regarding the intervention (highest score was 7 out of 10), implying that participants might have lacked an active role prior to

the start of intervention. Motivation is considered one of the most important factors for the success of a self-management programme [33,34]. Therefore, to increase the effectiveness of a non-pharmacological, multidisciplinary intervention in patients with GOA, attention should be paid to participants' motivation prior to inclusion. Furthermore, therapists could be trained in motivating and goal-setting techniques, for example motivational interviewing.

Several limitations should be taken into account when interpreting our data. First, we used a concurrent multiple-baseline single-case design to evaluate the intervention's potential effectiveness. This design is particularly successful in demonstrating immediate effects [35]. Since our intervention aimed to improve self-management in individuals with osteoarthritis, which is often considered challenging and time-consuming [9], our choice of study design might not be optimal, given the short evaluation period and the considerable length of the treatment programme. A second limitation was that all participants were in the same therapy group, possibly resulting in a negative group effect compromising any therapy effects. On the other hand, the traditional approach to multiple-baseline studies is for all participants to undergo treatment simultaneously [13]. This strategy is recommended as it improves internal validity, particularly in terms of history effects [36]. A third limitation, inherent to the design of the study, is that the study has lower external validity than randomised clinical trials, for which participants are usually selected to form a generalizable sample [37]. A fourth limitation of this study was its inability to test the feasibility of study logistics for a randomised clinical trial (for example, recruitment rate, drop-out rate, and issues concerning randomisation) [38]. A final limitation was that we selected patients based on their medical diagnosis and functional status rather than on their scores on our main secondary outcomes (i.e. pain and/or self-efficacy). Future studies should include clinically relevant thresholds for their outcome measures in the in- and exclusion criteria.

As far as we know, we are the first to study a multidisciplinary intervention to improve self-management in people with GOA. Due to differences in study populations, our results cannot be compared with those of another study into the effect of a non-pharmacological, multidisciplinary intervention in patients with GOA after major joint replacement surgery [39]. It is remarkable that so little research is available given the relatively high prevalence of individuals with established osteoarthritis with multiple joint involvement and its association with compromised health status [1,2].

Some consider single-case experimental designs as viable alternatives to large-scale randomised clinical trials [40,41], whereas others state the opposite [37,42]. Whilst using this design, we faced several (practical) constraints that potential users should be aware of. As yet, there is a plethora of analytical techniques for single-case data [31], with little or no consensus on the optimal way to analyse the data. In our study, we demonstrated a significant effect of our intervention on kinesiophobia using a randomisation test, whereas visual inspection showed only clear effects in one participant. Another practical

consideration is that the design requires a substantial contribution from the participants. In the present study, one of the participants dropped out as she experienced additional psychological burden due to recurring questionnaires. It remains to be elucidated whether frequent assessment of health status as in the current study negatively, or perhaps positively, influences health outcomes. In our opinion, the multiple-baseline single-case study is a useful and valid alternative to the randomised pilot study, as it gives insight into the feasibility of the intervention and allows to evaluate the intervention's potential effectiveness, allowing one to tailor the content and context of the intervention prior to conducting a randomised clinical trial. However, single-case studies should only be considered an alternative to a full-sized randomised clinical trial in rare diseases or in situations where a randomised clinical trial is unfeasible or unethical, because of the designs' limitations, including low external validity of the findings and the inability to correct for confounders (such as medication use, age, disease duration etc.).

An interesting finding was the marked variability in VAS scores within participants on specific outcomes. For example, three participants reported fluctuations in pain scores of more than 4 points within a period of half a week (i.e., between two measurement points). Fluctuations in pain between two measurement points ranged from 0 to 7 points, frequently exceeding the thresholds for clinically relevant differences [43]. Such fluctuations indicate that pain in OA is far less stable than often believed and should perhaps be assessed far more frequently. As such variations are also likely to occur in randomised clinical trials, researchers should consider assessing post-intervention health outcomes at repeated time points. These outcomes could then be averaged to obtain a more stable post-intervention point estimate.

In conclusion, health providers and researchers should be aware of the lack of studies on the effectiveness of non-pharmacological and/or multidisciplinary interventions for patients with GOA. In our study, although we systematically conceptualised our intervention according to the latest evidence [7-9] and in collaboration with several health care providers, both feasibility and effectiveness of the care programme are doubtful. Therefore, the therapy programme as described in this paper does not warrant evaluation in a large randomised clinical trial. Since the data obtained in this multiple baseline study have highlighted several ways in which the therapy program could be optimized/improved, these changes should be implemented prior to conducting an RCT to further examine the interventions' effectiveness.

REFERENCE LIST

- (1) Hoogbeem TJ, den Broeder AA, Swierstra BA, et al. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: A cross-sectional study. *Arthritis Care Res (Hoboken)* 2012;64:54-8.
- (2) Suri P, Morgenroth DC, Kwok CK, et al. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. *Arthritis Care Res (Hoboken)* 2010;62:1715-23.
- (3) Forestier R, Francon A, Briole V, et al. Prevalence of generalized osteoarthritis in a population with knee osteoarthritis. *Joint Bone Spine* 2011;78:275-8.
- (4) Gunther KP, Sturmer T, Sauerland S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. *Ann Rheum Dis* 1998;57:717-23.
- (5) Conaghan P, Birrell F, Burke M, et al. Osteoarthritis: national clinical guideline for care and management in adults. London: Royal College of Physicians 2008.
- (6) Stukstette M, Hoogbeem T, de RR, et al. A multidisciplinary and multidimensional intervention for patients with hand osteoarthritis. *Clin Rehabil* 2012;26:99-110.
- (7) Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthritis Cartilage* 2007;15:981-1000.
- (8) Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage* 2008;16:137-62.
- (9) Zhang W, Nuki G, Moskowitz RW, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage* 2010;18:476-99.
- (10) Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655.
- (11) Eldridge SM, Ashby D, Feder GS, et al. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clin Trials* 2004;1:80-90.
- (12) Backman CL, Harris SR, Chisholm JA, et al. Single-subject research in rehabilitation: a review of studies using AB, withdrawal, multiple baseline, and alternating treatments designs. *Arch Phys Med Rehabil* 1997;78:1145-53.
- (13) Christ T. Experimental control and threats to internal validity of concurrent and nonconcurrent multiple baseline designs. *Psychology in the Schools* 2007;44:451-9.
- (14) Fries JF, Spitz P, Kraines RG, et al. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137-45.
- (15) Hoogbeem TJ, Stukstette MJ, de Bie RA, et al. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. *BMC Musculoskelet Disord* 2010;11:142.
- (16) Nourbakhsh MR, Ottenbacher KJ. The statistical analysis of single-subject data: a comparative examination. *Phys Ther* 1994;74:768-76.

- (17) Holtgreffe K, McCloy C, Rome L. Changes associated with a quota-based approach on a walking program for individuals with fibromyalgia. *J Orthop Sports Phys Ther* 2007;37:717-24.
- (18) Ferron J, Sentovich C. Statistical Power of Randomization Tests Used With Multiple-Baseline Designs. *The Journal of Experimental Education* 2002;70:165-78.
- (19) de Jong JR, Vlaeyen JW, de Gelder JM, et al. Pain-related fear, perceived harmfulness of activities, and functional limitations in complex regional pain syndrome type I. *J Pain* 2011;12:1209-18.
- (20) Lorig K, Holman H. Arthritis Self-Efficacy Scales measure self-efficacy. *Arthritis Care Res* 1998;11:155-7.
- (21) Evers AW, Kraaimaat FW, van LW, et al. Beyond unfavorable thinking: the illness cognition questionnaire for chronic diseases. *J Consult Clin Psychol* 2001;69:1026-36.
- (22) Bijsterbosch J, Scharloo M, Visser AW, et al. Illness perceptions in patients with osteoarthritis: change over time and association with disability. *Arthritis Rheum* 2009;61:1054-61.
- (23) Beurskens AJ, de Vet HC, Koke AJ, et al. A patient-specific approach for measuring functional status in low back pain. *J Manipulative Physiol Ther* 1999;22:144-8.
- (24) Beurskens AJ, Bultmann U, Kant I, et al. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000;57:353-7.
- (25) Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. *J Psychol* 2005;139:439-57.
- (26) Maas M, Taal E, van der Linden S, et al. A review of instruments to assess illness representations in patients with rheumatic diseases. *Ann Rheum Dis* 2009;68:305-9.
- (27) Hoogboom TJ, den Broeder AA, de Bie RA, van den Ende CH. The Dutch Lower Extremity Functional Scale was highly reliable, valid and responsive in individuals with hip/knee osteoarthritis: a validation study. *BMC Musculoskeletal Disorders* 2012;13:117.
- (28) Beaton DE, Katz JN, Fossel AH, et al. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J Hand Ther* 2001;14:128-46.
- (29) Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. *Health Promot Int* 2006;21:245-55.
- (30) EpiData Data Entry, Data Management and basic Statistical Analysis System. Odense Denmark, EpidData Association, 2000-2008. [computer program]. 2012.
- (31) Bulte I, Onghena P. Randomization tests for multiple-baseline designs: an extension of the SCRT-R package. *Behav Res Methods* 2009;41:477-85.
- (32) Escobar A, Gonzalez M, Quintana JM, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. *Osteoarthritis Cartilage* 2012;20:87-92.
- (33) DeWalt DA, Davis TC, Wallace AS, et al. Goal setting in diabetes self-management: taking the baby steps to success. *Patient Educ Couns* 2009;77:218-23.
- (34) Teixeira PJ, Going SB, Houtkooper LB, et al. Exercise motivation, eating, and body image variables as predictors of weight control. *Med Sci Sports Exerc* 2006;38:179-88.

- (35) Neuman SB. Single-Subject Experimental Design. In: Duke NK, Mallette MH, editors. *Literacy Research Methodologies*. 2nd edition ed. New York: The Guilford Press; 2011. p. 383.
- (36) Carr JE. Recommendations for reporting multiple-baseline designs across participants. *Behav Intervent* 2005;20:219-24.
- (37) Newcombe RG. Should the single subject design be regarded as a valid alternative to the randomised controlled trial? *Postgrad Med J* 2005;81:546-7.
- (38) Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol* 2010;10:1.
- (39) Bulthuis Y, Drossaers-Bakker KW, Taal E, et al. Arthritis patients show long-term benefits from 3 weeks intensive exercise training directly following hospital discharge. *Rheumatology (Oxford)* 2007;46:1712-7.
- (40) Rizvi SL, Nock MK. Single-case experimental designs for the evaluation of treatments for self-injurious and suicidal behaviors. *Suicide Life Threat Behav* 2008;38:498-510.
- (41) Janosky JE. Use of the single subject design for practice based primary care research. *Postgrad Med J* 2005;81:549-51.
- (42) Guyatt GH, Keller JL, Jaeschke R, et al. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. *Ann Intern Med* 1990;112:293-9.
- (43) Lee JS, Hobden E, Stiell IG, et al. Clinically important change in the visual analog scale after adequate pain control. *Acad Emerg Med* 2003;10:1128-30.

10

Effectiveness of preoperative therapeutic exercise for major joint replacement on postoperative functional recovery: a meta-analysis and best practice evaluation.

ABSTRACT

Background. Our aim was to develop a rating scale to assess the therapeutic validity of therapeutic exercise. By use of this rating scale we investigated the therapeutic validity of therapeutic exercise in patients awaiting primary total joint replacement. Finally, we studied the association between therapeutic validity of preoperative therapeutic exercise and its effectiveness in terms of postoperative functional recovery.

Methods. (Quasi) randomised clinical trials on preoperative therapeutic exercise in adults awaiting joint replacement on postoperative recovery of functioning (self-reported or performance-based) during the hospital stay or within three months after surgery were identified through database and reference screening. Two reviewers extracted data and assessed the risk of bias and therapeutic validity. Therapeutic validity of the interventions was assessed with an nine-itemed expert-based rating scale (scores range from 0 to 9; score ≥ 6 reflecting therapeutic validity), developed in a four-round Delphi study. Effects were pooled using a random-effects model and meta-regression was used to study the influence of therapeutic validity.

Results. Of the 7492 articles retrieved, 12 studies (737 patients) were included. None of the included studies demonstrated therapeutic validity and two demonstrated low risk of bias. Therapeutic exercise was not associated with: 1) observed functional recovery during the hospital stay (Standardised Mean Difference [SMD]: -1.19; 95%-confidence interval [CI], -2.46 to 0.08); 2) observed recovery within three months of surgery (SMD: -0.15; 95%-CI, -0.42 to 0.12); and 3) self-reported recovery within three months of surgery (SMD -0.07; 95%-CI, -0.35 to 0.21) compared with control participants. Meta-regression showed no statistically significant relationship between therapeutic validity and pooled-effects.

Conclusion. Preoperative therapeutic exercise for major joint replacement did not demonstrate beneficial effects on postoperative functional recovery. However, poor therapeutic validity of the therapeutic exercise programmes may have hampered potentially beneficial effects, since none of the studies met the predetermined quality criteria. Future review studies on therapeutic exercise should address therapeutic validity.

Hoogeboom TJ, Oosting E, Vriezekolk JE, Veenhof C, Siemonsma PC, de Bie RA, van den Ende CHM, van Meeteren NLU.

Effectiveness of preoperative therapeutic exercise for major joint replacement on postoperative functional recovery: a meta-analysis and best practice evaluation.

PLoS ONE 2012;7(5),e38031.

INTRODUCTION

Total joint replacement is considered an effective and successful end-stage surgical procedure for relieving pain and improving functional status [1,2]. However, a significant number of patients experience persistent pain and functional disability after major joint replacement [3,4]. To enhance postoperative functional recovery, preoperative exercise is a potentially effective intervention by which to optimise the preoperative physical status of patients awaiting joint replacement [5,6]. However, systematic reviews are inconclusive regarding the effectiveness of preoperative exercise in terms of postoperative health status following total hip (THR) or total knee replacement (TKR) [7–10].

These reviews might be flawed as they fail to take into account the therapeutic validity of the exercise interventions in the individual studies, as recommended by Herbert and Bø [11]. It is known that, in the field of preoperative therapeutic exercise, there is a tendency for trials to include relatively healthy patients [12], rather than patients with known high-risk profiles for delayed postoperative recovery (patients of older age [13,14], with comorbidities and/or poor pre-operative status [13–19]), thus excluding patients for whom preoperative exercise is specifically indicated [20]. Furthermore, to yield optimal effects, the content of an exercise programme should be in line with the latest research, be of sufficient volume [21,22], and be tailored to the potential of the participants [23]. In terms of the latter, we hypothesize that poor therapeutic validity could result in negative study findings. To date, there is no clear set of criteria by which to assess the therapeutic validity of a therapeutic exercise intervention.

Therefore, the aim of our study was threefold. First, we developed a rating scale to assess the therapeutic validity of therapeutic exercise programmes. Second, we assessed the therapeutic validity of preoperative therapeutic exercise programmes in patients awaiting elective, primary THR or TKR, and, finally, we assessed the association between therapeutic validity and the effect of the interventions on postoperative functional recovery.

METHODS

The study comprised two phases: (1) a Delphi study to develop a rating scale for the therapeutic validity of therapeutic exercise, and (2) a systematic review and meta-analysis to assess the effectiveness of therapeutically valid exercise regimens in terms of observed functional recovery during the hospital stay, and in terms of self-reported and observed functioning after discharge within three months after surgery. This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [24,25].

Delphi rounds

For the Delphi rounds, we followed the method described by Yates et al. (2005) [26]. For the Delphi panel, we selected five, internationally renowned, Dutch experts on therapeutic exercise. All participants met the following criteria: (1) previous involvement in a published RCT of a therapeutic exercise treatment, (2) two or more published articles on therapeutic exercise, (3) two or more conference presentations on therapeutic exercise, and (4) licensed health professional in a relevant discipline. The experts were invited by e-mail to participate in the study. Anonymity among experts was maintained throughout all Delphi rounds.

The Delphi study was conducted over four rounds [26]. In the first round, participants responded to open-ended questions regarding therapeutic validity of therapeutic exercise. We defined therapeutic validity as '*the potential effectiveness of a specific intervention given the potential target group of patients*'. In the second round, the first and second authors collated and grouped the responses from round one into a number of statements regarding different aspects of therapeutic validity in therapeutic exercise. The expert group was then asked to determine which of the statements would be essential in a rating scale designed to measure the therapeutic validity of therapeutic exercise programmes (one point = very unnecessary, through to seven points = very necessary). In the third round, the first author created personalised questionnaires for each of the experts, comprising the median and inter-quartile range (IQR) of scores of each statement (representing group level of agreement and the degree of consensus, respectively) and the rating of the individual expert as a reminder. All experts then reviewed and re-rated the statements. A list of statements, which achieved consensus agreement, was prepared by the first author. Consensus for inclusion was defined as a median rating of six or seven on the seven-point rating scale and an IQR of 1.5 or less [26]. In the fourth and final round, all experts were allowed to anonymously express any final concerns regarding the list. These concerns were either accepted or declined by the whole expert group. Finally, the first and second authors drafted the output generated by the Delphi panel into a workable rating scale for the therapeutic validity of exercise programmes.

Systematic review

Search Strategy and Study Selection

We searched the following electronic databases (through to January 2012): MEDLINE (accessed by PubMed), Cochrane Central Register of Controlled Trials, EMBASE, ClinicalTrials.gov, CINAHL and PEDro. In addition, we manually searched the references of published studies. The initial search was not limited by language and comprised the terms *arthroplasty*, *exercise*, and related entry terms associated with a high-sensitivity

strategy for the search of RCTs [27]. The complete search strategies used for the different databases are shown in Appendix 2.

We included (quasi)RCTs that compared the effectiveness of preoperative structured therapeutic exercise training with a control intervention, with postoperative recovery of functioning (self-reported or performance-based) as an outcome in patients older than 18 years awaiting elective, primary THR or TKR. Structured exercise training was defined as an intervention in which patients were engaged in planned and supervised exercise programmes (i.e. resistance, aerobic or functional exercise). We only included studies that reported means, or differences between means, and respective dispersion values of postoperative functional recovery during the hospital stay and within 3 months after surgery. Exclusion criteria were: (1) duplicate publications or sub-studies of included trials, and (2) studies with two or fewer supervised exercise sessions. The comparator (control) group could be active (any non-exercise intervention) or placebo (no treatment or waiting list) group.

Titles and abstracts of retrieved articles were independently evaluated by two reviewers (TJH and JEV). Reviewers were not blinded to authors, institutions, or manuscript journals. Abstracts that did not provide enough information about the inclusion and exclusion criteria were retrieved for full-text evaluation. Reviewers independently evaluated full-text articles and determined eligibility for inclusion in review. Disagreements were resolved by consensus and, if disagreement persisted, by a third reviewer (C.H.M.E.). To avoid possible double counting of patients included in more than one report by the same authors or working groups, patient recruitment periods were evaluated and, if necessary, authors were contacted for clarification.

Data Extraction

Two reviewers (T.J.H. and E.O.) used standardised forms to independently extract the following information from each eligible publication: year of publication, geographical location, study population, functional outcome measures, duration of follow-up, and type and dose of exercise intervention. For the outcome measure of interest, the number of observations and means and standard deviations (SDs) were extracted for both the intervention and control groups at the following measurement points: 1) baseline (preoperative), 2) in-hospital (postoperative), and 3) after discharge (<3 months postoperative). If measures of variability were unavailable, we imputed the averaged SD of similar measures from other studies. If results were expressed as confidence intervals or interquartile ranges, we used transformation methods as recommended [28]. Where necessary, means and measures of dispersion were approximated from figures in the manuscripts using WebPlotDigitizer [29]. Characteristics of the exercise interventions were extracted, including the type, frequency, duration, and intensity. We used the Compendium of Physical Activities [21] to estimate the exercise intensity in terms of metabolic

equivalents (METs). Exercise volume (total energy expenditure on exercise, in METs·h¹·wk⁻¹) was calculated by multiplying the intensity in METs by total time spent exercising (number of exercise sessions multiplied by duration of each exercise session) [30].

Any disagreements about the extracted data were solved by consensus or by a third reviewer (C.H.M.E.). In case of missing data, the corresponding author of the included study was contacted.

Assessment of methodological (risk of bias) and therapeutic validity

Two reviewers (T.J.H and E.O.) independently assessed the methodological validity of the studies and the therapeutic validity of the therapeutic exercise programmes. The methodological validity (risk of bias) was scored using the adapted version of the Cochrane Collaboration's tool [31]. This adapted tool reviews five domains, with 11 items in total (see Appendix 1). Each item is rated as 'yes', 'no', or 'unsure'. Studies fulfilling six or more items were regarded as having a low risk of bias [32]. Therapeutic validity was scored using the rating scale developed in the Delphi rounds. Each item was rated as 'yes' or 'no'. Studies with six or more points out of nine were regarded as being of high therapeutic quality. Disagreements were resolved in a consensus meeting between the two raters. The strength of agreement between the two raters was measured by Cohen's κ coefficient (95%-confidence intervals), with $\kappa=0.41$ -0.60 indicating moderate agreement, $\kappa=0.61$ -0.80 representing good agreement, and $\kappa\geq 0.81$ representing very good agreement [33].

Data analysis

In this study, we compared structured, valid therapeutic exercise with a control intervention at three different outcome levels, namely 1) observed functional recovery during the hospital stay; 2) recovery of self-reported functioning within three months of surgery; and 3) recovery of observed functioning within three months of surgery. In our primary analyses, we only included highly valid studies (i.e. risk of bias score > 6 & therapeutic validity score > 5). Sensitivity analyses were performed without any restrictions on validity. All analyses were carried out separately for patients awaiting either TKR or THR. When more than one study was available, data were statistically pooled where appropriate.

Measures of functioning (performance and self-reported measures) in the treatment and control groups were transformed to standardised mean differences (Hedges g) to cope with the variety of outcome measures [28,34]. To ensure uniform interpretability of all scales (i.e., higher scores representing more functional problems), we transformed our data according to the Cochrane recommendations [28]. For studies that compared multiple exercise interventions with a single control group, we split this shared control group into two or more subgroups with smaller sample sizes weighted in relation to different exercise interventions. We applied this approach to ensure reasonably independent comparisons and to overcome a unit-of-analysis error for studies that could contribute to multiple and

correlated comparisons [28]. Calculations were performed using a random-effects model. An α value of < 0.05 was considered statistically significant.

We assessed statistical heterogeneity of the treatment effect among studies using the inconsistency I^2 test, in which values greater than 50% were considered indicative of high heterogeneity [28]. To assess heterogeneity between studies, we reran the meta-analyses whilst removing one study at a time to check if a particular study caused heterogeneity.

To explore whether effects of the exercise interventions on functional recovery were associated with therapeutic validity (0-9 points) or by exercise volume ($\text{METs} \cdot \text{h}^{-1} \cdot \text{wk}^{-1}$), we performed meta-regression analyses on each of the three outcome points (i.e. in-hospital functional recovery, short-term observed functional recovery, and short-term self-reported functional recovery), whilst accounting for hip or knee replacement. We evaluated the goodness of fit of each model using the adjusted R^2 , which denotes the proportion of between-study variation explained by the covariates.

Publication bias was assessed using a contour-enhanced funnel plot of each trial's effect size against the standard error [35]. Funnel plot asymmetry was evaluated by Begg and Egger tests, and a significant publication bias was considered to be present if the P value was less than 0.10. If publication bias was apparent, trim-and-fill computation was used to estimate the effect of publication bias on the interpretation of results [35,36].

All analyses were conducted using Stata software, version 10.0 (Stata Inc., College Station, Texas).

RESULTS

Delphi study

Delphi study

The initial open-ended questionnaire was sent to five experts in the field of therapeutic exercise, all of whom met our predetermined criteria. All five experts responded to the invitation and completed each of the four Delphi rounds; no attrition occurred. The experts agreed unanimously that trials on exercise therapy should be assessed on therapeutic validity and that therapeutic validity should be accounted for in best evidence synthesis in systematic reviews.

After the first round, a total of 49 unique statements were generated which could be aggregated into 10 recurrent themes (see Appendix 3). After the second round, consensus was reached on 22 out of the 49 statements (45%). The highest level of disagreement (i.e. largest IQR) was found for the item: "The exercise programme is personalised for each participant". The lowest score was found for the item: "Natural fluctuations in disease activity must be controlled for." In the third round, full consensus (i.e. median = 7 and IQR

= 0) was not reached for any of the items, although for 10 items the degree of consensus was zero with a median score of six. In the fourth and final round, eight concerns were expressed regarding the pre-final list, mostly due to item formulation (n = 4).

In the final phase, the expert panel considered the 22 statements generated by the Delphi panel and collated them into a nine-item rating scale covering five critical areas. This scale was named the CONTENT (Consensus on Therapeutic Exercise Training) scale (see Table 1).

Table 1. The CONTENT scale for the therapeutic validity of therapeutic exercise programmes.

Items		Judgement	
A. Patient eligibility			
1.	Was the patient selection described? To score “yes”, patient selection should be described and participants should be screened for contraindications (for instance, using red and yellow flags) (this must be explicitly mentioned in the manuscript; otherwise “no”).	Yes	No
2.	Was the patient selection adequate? This item can be scored as “yes” if: <ul style="list-style-type: none">the goals of the therapeutic exercise match the participants’ problems (for instance, if the goal of the therapeutic exercise is to improve a patients’ functional status, then only patients with deprived functional status should be included). In this case participants’ problems represent bodily functions and structures, activities and participation levels, see the 'International Classification of Functioning, Disability and Health (ICF); andthe selection criteria match the majority of potential participants. Ergo, the therapeutic exercise should not be evaluated in a population that - in clinical practice - is nearly non-existent.	Yes	No
B. Competences and setting			
3.	Were eligibility criteria for therapist and setting determined and adequate? The questions to be answered here are: <ul style="list-style-type: none">Are the goals and content of the therapeutic exercise matched to the therapist’s competences and skills?Are the goals and content of the therapeutic exercise matched to the location or setting where the therapeutic exercise takes place? If no eligibility criteria are described, this item should be scored as “no”.	Yes	No
C. Rationale			
4.	Was the therapeutic exercise based on a-priori aims and intentions? Did the authors describe a-priori aims, intentions and hypotheses about the therapeutic exercise on theoretically driven and/or argued choices? If this question can be answered with “yes”, this item is scored as “yes”.	Yes	No
5.	Was the rationale for the content and intensity of the therapeutic exercise described and plausible? Did the authors describe why they believed the content (e.g. resistance exercise training, aerobic exercise training, flexibility training, etc.) and intensity (e.g. moderate/vigorous intensity, length of exercise, etc.) of the studied intervention was likely to achieve their treatment goals?	Yes	No

D. Content

- | | | | |
|----|--|-----|----|
| 6. | Was the intensity of the therapeutic exercise described?
This item can be scored as “yes” if: | Yes | No |
| | <ul style="list-style-type: none"> ○ the content of the therapeutic exercise is described in specific terms (i.e. duration, frequency and intensity of exercise sessions (e.g. 80% $\text{VO}_{2\text{max}}$, level of exertion (RPE), repetition maximum, etc.) and the total duration of the therapeutic exercise); ○ the intensity of the therapeutic exercise was selected and adjusted on theoretically driven and/or argued choices; and ○ the content of the therapeutic intervention is suitable for the majority of patients.. | | |
| 7. | Was the therapeutic exercise monitored and adjusted when considered necessary?
This item can be scored as “yes” if: | Yes | No |
| | <ol style="list-style-type: none"> 1. the regular and structured monitoring of therapy progression allows the therapist to: <ul style="list-style-type: none"> ○ strive for optimal exercise intensity; ○ adjust the intervention in case of therapy failure on an individual level; and ○ identify and monitor adverse events. 2. the outcome measures match the therapy goals. | | |
| 8. | Was the therapeutic exercise personalised and contextualised to the individual participants?
The goals and content of the therapeutic exercise should not only match the patients’ bodily functions and structures, activities and participation levels, but also their personal and environmental factors (see ICF). This item can be scored as “yes” if the therapeutic exercise accounts for relevant personal (e.g. motivation, coping, ethnicity, etc.) and environmental (e.g. logistics, support family/friends, products and technology, etc.) factors for each of the included participants. | Yes | No |

E. Adherence

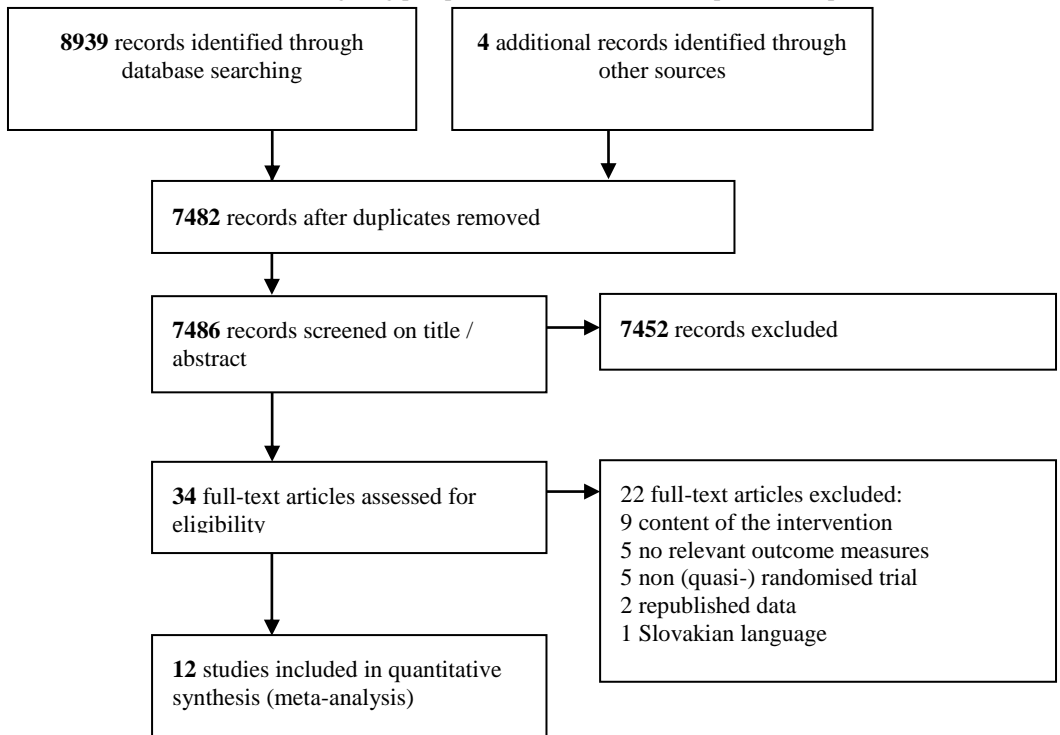
- | | | | |
|----|--|-----|----|
| 9. | Was adherence to the therapeutic exercise determined and acceptable?
For adherence to be properly described and acceptable, adherence should be described in such a way that it allows the reader to understand whether the actual executed therapeutic exercise differed from the planned therapeutic exercise (i.e. data should be provided on the achieved intensity, for example number of sessions attended, achieved exercise intensity, number of exercises etc.). Moreover, adherence should be quantitatively known, allowing it to be controlled for in the analysis. | Yes | No |
|----|--|-----|----|

*Systematic review***Description of studies**

We identified a total of 8939 records in the initial search and removed 1457 duplicate publications. We excluded 7452 non-relevant records based on title or abstract screening. Full-text articles were retrieved for 34 publications and assessed for eligibility. Twelve English-language articles comprising 11 randomised controlled trials and one quasi-randomised controlled trial met the eligibility criteria [37–48] (Figure 1).

One study presented data for both THR and TKR [45], therefore eight interventions on TKR and five interventions on THR were included. Moreover, one TKR study presented data for 2 comparisons [38], resulting in nine interventions in the TKR group. These 12 studies included a total of 737 patients (55% women), with a mean (SD) age of 66 (8) years and a Body Mass Index (BMI) of 31 (6).

Figure 1. Selection of trials investigating preoperative exercise for total hip or knee replacement.



The therapeutic exercise interventions prior to TKR and THR are described in Tables 2 and 3, respectively. Of the eight studies ($n = 502$) on therapeutic exercise prior to TKR, eight investigated resistance exercise [37–39,44–48] and one investigated aerobic exercise [38]. Typically, these interventions were carried out 3 times a week for 5 weeks, at an intensity of $7.2 \text{ METs} \cdot \text{h}^{-1} \cdot \text{wk}^{-1}$ (see Table 2). Of the five studies ($n = 235$) on therapeutic exercise prior to THR, four studied resistance exercise [40–42,45] and one examined functional exercise [43]. Typically, these interventions were carried out 2.5 times a week for a period of 6 weeks and at an intensity of $10.9 \text{ METs} \cdot \text{h}^{-1} \cdot \text{wk}^{-1}$ (see Table 3).

Risk of Bias and Publication Bias assessment

Appendix 4 shows the methodological quality assessment of individual studies. The initial agreement of the reviewers on the total risk of bias assessment was 85% (112 of 132 items), and Cohen's Kappa (95%-CI) was 0.77 (0.67 - 0.85). All disagreements were resolved in a consensus meeting. Ten studies were assessed as having a high risk of bias and two studies were assessed as having a low risk of bias [37,43]. The most prevalent limitations were found in items about blinding (patient, care provider, outcome assessor), allocation concealment, compliance and intention-to-treat analysis.

For the in-hospital recovery data, the Egger regression test suggested funnel plot asymmetry ($P = 0.07$), indicating publication bias. After applying the trim-and-fill procedure, we estimated that two studies were missing, and the adjusted estimate of overall SMD was -2.43 (95% CI, -3.77 to -1.08, $P < 0.01$). Contour-enhanced funnel plots and statistical tests did not show any publication bias for the short-term post-operative observational data (Egger: $P = 0.41$ and Begg $P = 0.54$) and the self-reported data (Egger: $P = 0.47$ and Begg: $P = 0.18$).

Table 2. Description of Supervised Exercise Intervention for patients awaiting Total Knee

Source, y	Study Location	Age, y [†]	Women, %	BMI [†]	Exercise group		Control group		Trial design
					No of sub-jects	Exercise intervention	No of sub-jects	Control intervention	
Beaupre et al, 2004	Canada	67 (6.5)	55	31.5 (5.5)	65	Resistance exercise	66	No intervention	RCT
D'Lima et al, 1996	USA	69 (5.5)	60	NA	10	Resistance exercise	10	Education session and leaflet	RCT
Evgeniadis et al, 2008	Greece	68.3 (3.5)	61	34.1 (5.0)	24	Resistance exercise	24	No intervention	RCT
Rodgers et al, 1998	USA	67.6 (18.4)	55	NA	12	Resistance exercise	11	No intervention	Quasi RCT
Rooks et al, 2006	USA	67.0 (8.2)	54	34.8 (7.9)	22	Resistance exercise	23	Education via leaflet and telephone	RCT
Topp et al, 2009	USA	63.8 (6.8)	NA	32.1 (5.9)	26	Resistance exercise	28	No intervention	RCT
Weidenhielm et al, 1993	Sweden	63.5 (4.5)	52	29.6 (0.5)	20	Resistance exercise	20	No intervention	RCT
Williamson et al, 2007	UK	69.8 (9.4)	53	32.7 (6.1)	60	Resistance exercise	61	Education leaflet	RCT
D'Lima et al, 1996	USA	70.6 (6.5)	35	NA	10	Aerobic exercise	10	Education session and leaflet	RCT
Total		66 (8)	54	32 (5)	249		253		8

Abbreviations: AT: Aerobic Training, BMI: Body Mass Index, MET: Metabolic Equivalent, OT: Occupational

[†]Values are expressed as mean (SD). [‡]Amount of energy expenditure per week during programmed exercise

Replacement (continuation Table 2).

Deliverer	Type	Supervised exercise intervention			MET, h/wk [‡]
		No. of times/wk	No. of weeks	Intensity	
PT	Cycling, lower extremity weight training	3	4	AT: 7.5 minutes at low intensity. WT: 3 sets, 10-15 repetitions, 5 exercises, intensity unclear (progressively increased to patients' tolerance).	6.6
PT	Lower & Upper extremity weight training	3	6	WT: 45 minutes at intensity tolerated by patient (adjusted with one repetition every third day).	6.5
PT + OS	Core & Upper extremity weight training	3	4	WT: 3-4 sets, 10-14 repetitions, ? exercises, intensity based on patient's ability to perform 1 set of 8 repetitions (progressed when 15 repetitions are performed comfortably).	5.3
PT	Cycling, lower extremity weight training	3	6	AT: ? minutes at intensity according to baseline capacity. WT: ? sets, ? repetitions, 6 exercises, intensity unclear (adjusted after 3 weeks).	?
PT	Cycling, total body weight training (3 wks), aquatic training (3 wks)	3	6	AT: 10 minutes at moderate intensity. WT: 2 sets, 10 repetitions, 7 exercises, intensity tailored to person's fitness level and comfort with exercises (progression unclear).	10.0
RES	Lower extremity weight training	3	4.3	WT: 2 sets, 10 repetitions, 9 exercises, intensity low (adjusted by increasing sets/resistance)	7.6
PT	Cycling, lower extremity weight training	3	5	AT: 10 minutes at 50 turns/min, unloaded. WT: 2 sets, 10 repetitions, 5 exercises, intensity against gravity to 3kg (progression 10RM principle).	6.4
PT	Lower extremity weight training	1	6	WT: 60 minutes, ? sets, ? repetitions, 9 exercises, intensity unclear (progression unclear).	4.3
PT	Cycling, arm ergometry, aquatic training	3	6	AT: 45 minutes at a heart rate of Resting Heart Rate + 0.4-0.7·Heart Rate Reserve.	10.7
		2.8 (0.4)	5.3 (0.8)		7.2 (1.7)

Therapist, PT: Physical Therapist, RCT: Randomised Clinical Trial, RES: Researcher, WT: Weight training.
(1 metabolic equivalent equals 1 kcal·kg⁻¹·hour⁻¹).

Table 3. Description of supervised exercise intervention for patients awaiting Total Hip

Source, y	Study Location	Age, y [†]	Women, %	BMI [†]	Exercise group		Control group		Trial design
					No of sub-jects	Exercise intervention	No of sub-jects	Control intervention	
Ferrara et al, 2008	Italy	63.4 (7.8)	61	NA	11	Resistance exercise	12	No intervention	RCT
Gilbey et al, 2003	Australia	65.2 (11.1)	65	27.9 (4.3)	37	Resistance exercise	31	No intervention	RCT
Rooks et al, 2006	USA	62.0 (9.7)	58	29.3 (7.4)	32	Resistance exercise	31	Education via telephone + leaflet	RCT
Gocen et al, 2004	Turkey	51.3 (13.6)	36	26.3 (3.9)	30	Resistance exercise	30	No intervention	RCT
Hoogeboom et al, 2010	the Netherlands	76.0 (4.2)	67	31.6 (11.3)	10	Functional exercise	11	One education session	Pilot RCT
Total		62 (10)	56	28 (6)	120		115		5

Abbreviations: ACSM: American College of Sports Medicine, AT: Aerobic Training, BMI: Body Mass Index, Trial, WT: Weight training. [†]Values are expressed as mean (SD). ^{*}Amount of energy expenditure per week during

Replacement (continuation Table 3).

Supervised exercise intervention					
Deliverer	Type	No. of times/wk	No. Of weeks	Intensity	MET, h/wk [‡]
PT	Cycling, lower extremity weight training	5	4	AT: 12.5 minutes at low/moderate intensity. WT: 3-4 sets, 8-12 repetitions, ? exercises, intensity unclear (progression unclear).	16.0
PT	Cycling / arm ergometry / rowing, + aquatic training	2	8	AT: 5 min cycling/arm/rowing (warming up), 10 min aquatic cycling/running. WT: 1-3 sets, 10 repetitions, ? exercises. All intensity according to ACSM guidelines (sets increased when patient improved).	7.4
PT	Cycling, total body weight training (3 wks), aquatic training (3 wks)	3	6	AT: 10 minutes at moderate intensity. WT: 2 sets, 10 repetitions, 7 exercises, intensity tailored to person's fitness level and comfort with exercises (progression unclear).	10.0
PT	Upper extremity weight training	0.5	8	Three times a day over an eight week period, patients performed 10 repetitions of ? exercises. A PT visited the patients at a two week interval.	?
PT	Walking, cycling, lower extremity weight training, functional training	2	4.5	AT: 5 min walk (warming up), 25 min cycling. WT: 1 set, 15 repetitions, 2 exercises. FT: 3 sets, 15 repetitions, 10 exercises. Intensity for all exercises was moderate to high (13-14 RPE) and adjusted when patient's rated an exercise 12 RPE.	10.2
		2.5 (1.2)	6.1 (1.5)		10.9 (2.6)

MET: Metabolic Equivalent, OT: Occupational Therapist, PT: Physical Therapist, RCT: Randomised Clinical programmed exercise (1 metabolic equivalent equals 1 kcal·kg⁻¹·hour⁻¹).

Table 4. Methodological and therapeutic validity scores per study.

Study	Methodological Validity (0-11)	Therapeutic Validity (0-9)
Beaupre et al (2004)	7 (64%)	1 (11%)
D'Lima et al (1996) (RE)	3 (27%)	1 (11%)
D'Lima et al (1996) (AE)	3 (27%)	2 (22%)
Evgeniadis et al (2008)	4 (36%)	2 (22%)
Ferrara et al (2008)	5 (45%)	0 (0%)
Gilbey et al (2003)	2 (18%)	1 (11%)
Gocen et al (2004)	3 (27%)	0 (0%)
Hoogeboom et al (2010)	7 (64%)	5 (56%)
Rodgers et al (1998)	2 (18%)	2 (22%)
Rooks et al (2006)	4 (36%)	3 (33%)
Topp et al (2009)	3 (27%)	2 (22%)
Weidenhielm et al (1993)	4 (36%)	0 (0%)
Williamson et al (2007)	4 (36%)	1 (11%)

Abbreviations: AE = Aerobic exercise, RE = Resistance exercise.

Therapeutic validity assessment

Appendix 5 shows the therapeutic validity assessment score per individual study as assessed using the CONTENT scale. Cohen's kappa revealed a moderate agreement between the two raters of 0.70 (0.62 – 0.78); absolute agreement was 104 out of 117 items (89%). The item "Was the therapeutic exercise based on a-priori aims and intentions?" had the least agreement between the raters. All disagreements were resolved without consulting the third rater. The median score (IQR) and mean score (range) of the therapeutic quality of interventions was 1 (1) and 1.5 (0-5), respectively. None of the 13 interventions could be labelled as being therapeutically valid according to the cut-off score of six or higher. Both therapeutic validity and methodological validity scores are presented in Table 4.

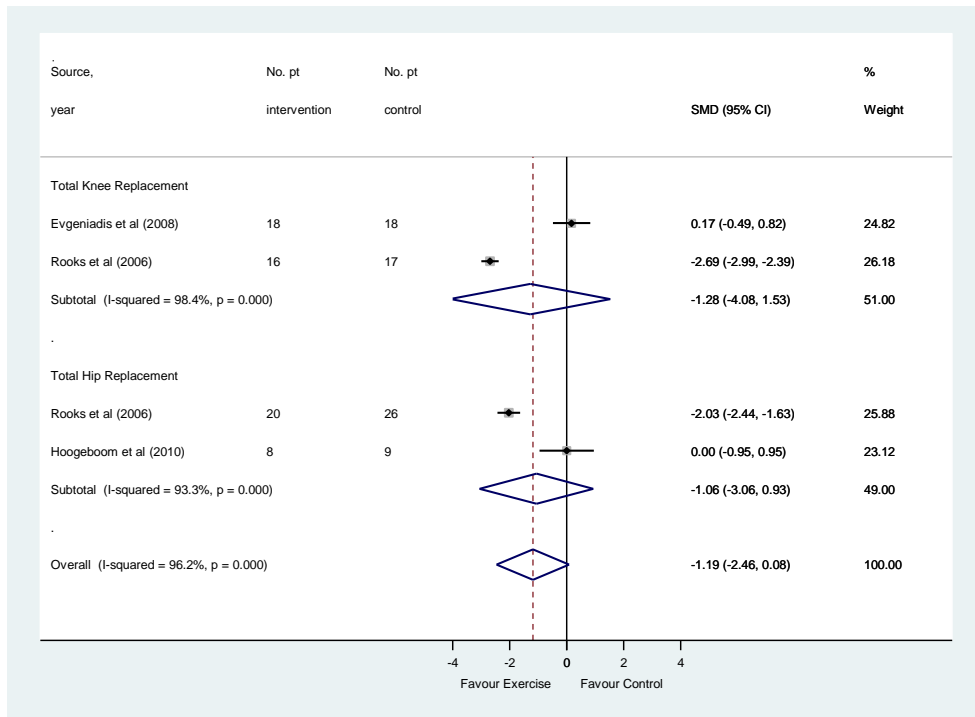
The categories 'Setting and Therapist', 'Monitoring', and 'Adherence' had the lowest score; none of the interventions included these aspects in their intervention. The highest-scoring category was 'Rationale of the study', with nine out of 13 studies scoring 'Yes' (69%). Two studies (15%) provided a rationale for the content of the therapy. Patient selection was described in four interventions (31%), but only one intervention (8%) was in line with the described aims and intentions of the intervention. Intensity of the intervention was described adequately in three of the 13 interventions (23%).

Association between intervention and in-hospital functional recovery

None of the three studies (132 patients) in this category met the requirements for methodological and therapeutic validity [39,43,45]. Sensitivity analysis of the overall

pooled effect of structured preoperative exercise vs. control in terms of functional recovery during the hospital stay was -1.19 (95% CI, -2.46 to 0.08 ; I^2 , 96.2%; P for heterogeneity <0.001) (Figure 2).

Figure 2. Functional recovery during hospital stay in individual studies of structured exercise training vs. control intervention.



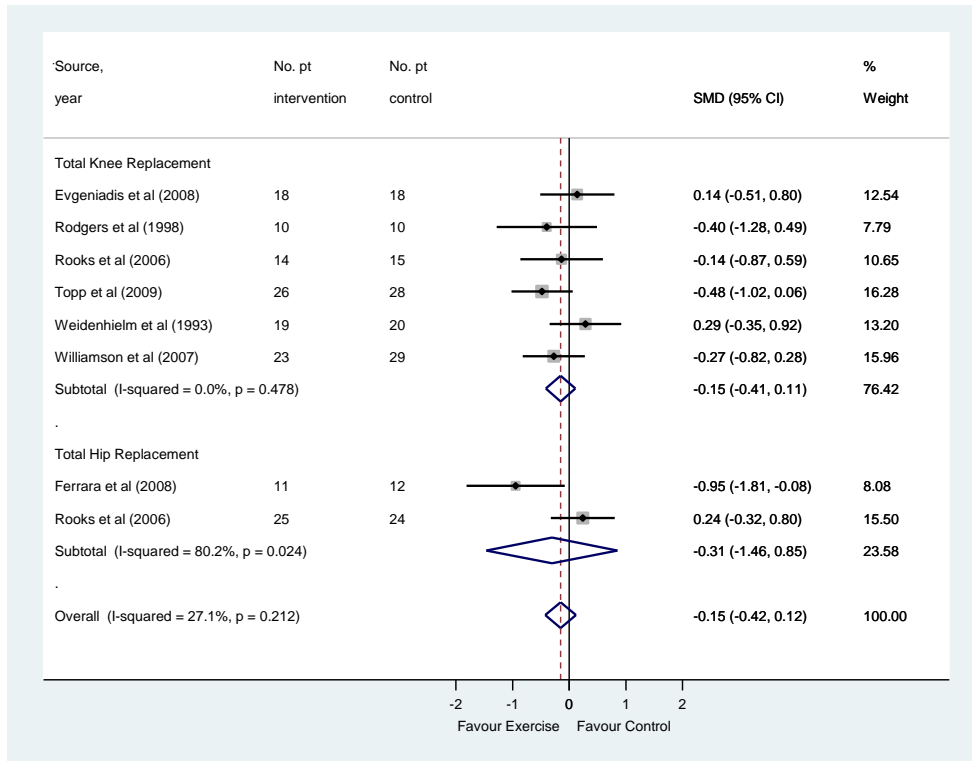
Similar pool effects were found when the analysis was separated into THR [43,45] and TKR [40,45], albeit with broader 95% confidence intervals (Figure 2). Meta-regression did not demonstrate an association between the pooled effect and exercise volume ($\beta = -1.70$; 95%-CI $-21.56 - 18.15$) or therapeutic validity score ($\beta = 0.32$; 95%-CI $-13.23 - 13.87$).

Association between intervention and short-term observed functional recovery

None of the seven studies in this category met the requirements for methodological or therapeutic validity [39,40,44–48]. Disregarding any predetermined validity scores, sensitivity analyses found that overall short-term observed functional status was not

associated with structured exercise; SMD -0.15 (95% CI, -0.42 to 0.12; I^2 , 27.1%, P for heterogeneity = 0.212) (Figure 3).

Figure 3. Short-term recovery of observed functioning in individual studies of structured exercise vs. control intervention.

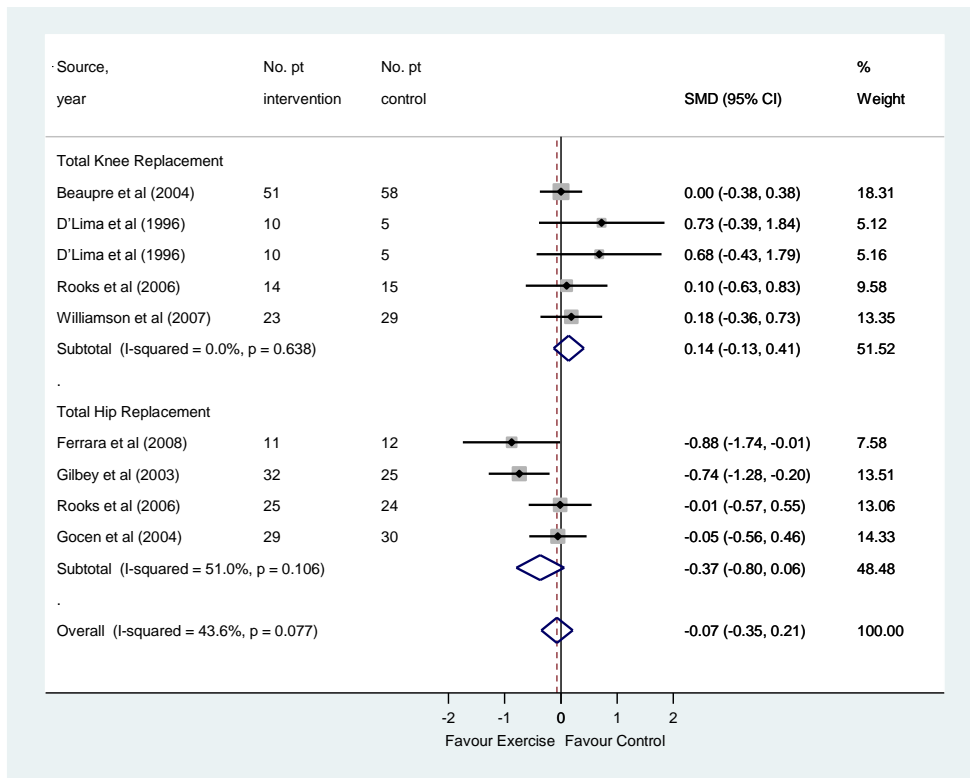


For the TKR subgroup (6 studies, 230 patients) [39,44–48], random-effect modelling revealed a non-significant SMD for the effect of structured exercise on observed functional recovery, SMD -0.15 (95% CI, -0.41 to 0.11; I^2 , 0.0%, P for heterogeneity = 0.478). For the THR subgroup (2 studies, 72 patients) [40,45], a non-significant SMD of -0.31 (95% CI, -1.46 to 0.85, I^2 , 80.2%, P for heterogeneity = 0.024) was found for the effect of structured preoperative exercise on observed functional recovery. Meta-regression demonstrated no association between the interventions' short-term effects on functional recovery and exercise volume (β = -0.15; 95%-CI -0.364 - 0.07) or therapeutic validity (β = 0.08; 95%-CI -0.09 - 0.26).

Association between intervention and short-term self-reported functional recovery

Methodological validity was demonstrated in one of the seven studies in this category [37], while therapeutic validity was found in none. Sensitivity analysis of the seven studies comparing structured exercise (205 patients) vs. control (203 patients)[37,38,40–42,45,48], showed that exercise was not associated with self-reported short-term functional recovery after major joint replacement; SMD -0.07 (95% CI, -0.35 to 0.21; I^2 , 43.6%, P for heterogeneity = 0.077) (Figure 4).

Figure 4. Short-term recovery of self-reported functioning in individual studies of structured exercise vs. control intervention.



For the TKR subgroup [37,38,45,48], the overall association between five structured therapeutic exercise programmes vs. control and short-term self-reported functioning was 0.14 (95% CI, -0.13 to 0.41; I^2 , 0.0%, P for heterogeneity = 0.638). For the THR subgroup [40–42,45], random-effect models of four studies (188 patients) on structured exercise revealed a non-significant SMD in favour of structured exercise; SMD -0.37 (95% CI, -0.80

to 0.06; I², 51.0%, *P* for heterogeneity = 0.106). Meta-regression showed no association between pooled effects and exercise volume ($\beta = 0.02$; 95%-CI -0.15 - 0.19)) or therapeutic validity ($\beta = -0.01$; 95%-CI -0.18 - 0.15)).

DISCUSSION

Our results demonstrate that the effectiveness of (highly) valid, structured therapeutic exercise training in individuals awaiting major joint replacement surgery remains unconfirmed. Of the 12 eligible studies, only two met the requirements for methodological quality and none met the prespecified requirements for therapeutic validity, highlighting a lack of quality in this field. Furthermore, pooling data from all eligible studies showed no benefit of preoperative therapeutic exercise therapy in terms of functional recovery after THR or TKR. These findings should, however, be interpreted with caution.

Expert opinion in our Delphi rounds identified five critical areas, comprising a total of 9 items, as being important for the therapeutic validity of a therapeutic exercise intervention. These five critical areas are patient selection, therapist and setting selection, rationale, content, and adherence, and are supported by evidence from the literature. For example, several studies have demonstrated that adequate patient selection can be of great importance in treatment effectiveness, as some patients respond differently to non-pharmacological interventions than others [49–51]. Thus, proper patient selection might result in greater therapy gains [52]. In addition, the selection of therapist and setting are also both known to influence treatment effects [53]. Furthermore, a plausible rationale regarding the benefits of the therapeutic exercise programme - especially if there is little or no previous experience with the intervention - is thought to be necessary to achieve therapy effects [54]. In fact, studies lacking a clear rationale are even considered to be unethical [55]. Adequate intervention content, characterised by sufficient dosing based on theoretical or argued choices, monitoring and personalisation, is perhaps the most important factor in yielding therapy effects. For example, evidence shows that strength training programmes produce the greatest increases in muscle strength if the training load is high [22] without the consideration of frailty [56]. The use of intermediate outcomes is also essential to optimally dose the therapeutic exercise intervention, to achieve therapy progress, and to prevent therapy failure [57]. Finally, the last critical area identified by the Delphi group was adherence to the intervention. Adherence to the exercise programme determines the extent to which therapy dosing is indeed achieved [58]. Therefore, it has been recommended that exercise programmes should be described in sufficient detail to enable readers to understand how the intervention was actually carried out [11]. In conclusion, each of the five aspects of therapeutic validity identified by the Delphi study is supported by the literature.

Our finding that preoperative therapeutic exercise has no beneficial effect on functional recovery after joint replacement surgery is in line with our hypothesis that suboptimal therapeutic exercise elicits no effect. None of the included studies met the predetermined requirements for therapeutic validity. An apposite example demonstrating this lack of therapeutic validity is that, although nine out of 13 exercise interventions provided a rationale for why preoperative exercise would elicit beneficial effects, only one group [43] actually applied their rationale to their patient selection criteria (i.e. by including patients with a high risk of delayed functional recovery), and only two studies [43,45] applied this rationale to their exercise programme (i.e. by selecting their exercise dosing accordingly). Moreover, none of the included interventions monitored therapy dosing to achieve and maintain optimal exercise dosing [57], as is further illustrated by the finding that only three studies [38,40,43] reported a supervised exercise dose greater than the regularly prescribed weekly amount of physical activity (i.e. $10 \text{ METs} \cdot \text{h}^{-1} \cdot \text{wk}^{-1}$) [59]. Finally, adherence was often not, or only marginally, reported. Apart from the number of attended sessions, authors should provide information on the prespecified exercise protocol and whether the intended exercise intensity was reached. In conclusion, we recommend that future studies on preoperative therapeutic exercise develop a highly valid therapy protocol, for which our rating scale could be used as a blueprint.

For an exercise programme to be considered therapeutically valid, we arbitrarily chose a cut-off value of six out of nine items on the CONTENT scale. Lowering the cut-off score to five or even four points would not have altered our conclusions regarding short-term postoperative functional recovery. Regarding the in-hospital functional recovery, lowering the cut-off score to four or five would have identified one pilot trial [43] that was insufficiently powered to assess differences in postoperative recovery. Whether the current cut-off value represents a true threshold for therapeutic validity needs to be further investigated.

Risk of bias

Ten out of 12 studies were considered to have a high risk of bias. Allocation concealment and blinding were the lowest scoring items in the risk of bias assessment. Because most of the studies lack allocation concealment, readers should be aware that these studies are more susceptible to selection bias, and this may affect the generalisability of our results. Moreover, given that most studies were insufficiently blinded and that the majority of studies did not use intention-to-treat analysis, the apparent results of our meta-analysis may have been inflated [60,61].

Limitations

Since effectiveness in randomised trials depends on the quality of the intervention, the lack of criteria to assess this quality is surprising. To date, some systematic reviews have investigated the relationship between exercise intensity and therapeutic effectiveness post-

hoc [30,62], with varying effects. One limitation of our study is that we were unable to draw conclusions regarding the validity of our rating scale, as none of the included studies could be classified as being highly valid. In fact, the majority of the interventions scored in the lowest tertile of the scale, preventing us from evaluating the relationship between therapy outcomes and therapy validity. Another limitation is that the CONTENT-scale might not only evaluate the therapeutic validity of an exercise program but also how well the exercise program was justified and how completely the justification was reported. Perhaps some of the studies employed adequate exercise programs but scored poorly on the scale because the study reports did not include a complete justification of the exercise programs.

Comparison with existing literature

So far, several systematic reviews [7,8,63], narrative reviews [9,64], and meta-analyses [10,65] have been published on preoperative exercise in patients awaiting joint replacement, but none of these reviews assessed the quality of the included interventions [11]. Taken the therapeutic validity into account, we have reached a similar conclusion to previous reviews, namely that the current intervention studies, which is mainly of low methodological validity, does not show that therapeutic exercise has beneficial effects on postoperative outcomes. However, what our review adds is that readers should also take the low therapeutic validity into consideration when interpreting these conclusions. Future studies should therefore specifically aim to include patients at need, that is those at risk for postoperative delayed recovery (based on a validated clinical decision rule) [52], provide a (piloted) [23] therapeutically sound and feasible exercise programme of sufficient, titrated dosing [57] and evaluated on relevant and amendable parameters (for instance heart rate recovery) [66]. The preoperative exercise program for patients awaiting coronary artery bypass grafting reported by Hulzebos et al (2006) is an illustration of the systematic development of an exercise program while addressing critical areas for therapeutic validity [20].

Conclusions

In conclusion, none of the 13 included therapeutic exercise programmes met our predetermined criteria for high therapeutic validity, making it unlikely that the interventions evaluated in these studies would have elicited relevant effects. In our view, the interpretation and development of therapeutic exercise programmes would be facilitated if international consensus could be reached on a select number of mandatory criteria for therapeutic validity. Finally, we recommended that future review studies on therapeutic exercise should not only determine the methodological validity, but also the therapeutic validity of the included trials.

REFERENCE LIST

- (1) Ewald FC, Wright RJ, Poss R, Thomas WH, Mason MD, Sledge CB (1999) Kinematic total knee arthroplasty: a 10- to 14-year prospective follow-up review. *J Arthroplasty* 14: 473-480.
- (2) Anderson JG, Wixson RL, Tsai D, Stulberg SD, Chang RW (1996) Functional outcome and patient satisfaction in total knee patients over the age of 75. *J Arthroplasty* 11: 831-840.
- (3) Hawker GA (2006) Who, when, and why total joint replacement surgery? The patient's perspective. *Curr Opin Rheumatol* 18: 526-530.
- (4) Wylde V, Dieppe P, Hewlett S, Learmonth ID (2007) Total knee replacement: is it really an effective procedure for all? *Knee* 14: 417-423.
- (5) Ditmyer MM, Topp R, Pifer M (2002) Prehabilitation in preparation for orthopaedic surgery. *Orthop Nurs* 21: 43-51.
- (6) Hoogbeem TJ, van den Ende CH, van der Sluis G, Elings J, Dronkers JJ, Aiken AB, van Meeteren NL (2009) The impact of waiting for total joint replacement on pain and functional status: a systematic review. *Osteoarthritis Cartilage* 17: 1420-1427.
- (7) Ackerman IN, Bennell KL (2004) Does pre-operative physiotherapy improve outcomes from lower limb joint replacement surgery? A systematic review. *Aust J Physiother* 50: 25-30.
- (8) Coudeyre E, Jardin C, Givron P, Ribinik P, Revel M, Rannou F (2007) Could preoperative rehabilitation modify postoperative outcomes after total hip and knee arthroplasty? Elaboration of French clinical practice guidelines. *Ann Readapt Med Phys* 50: 189-197.
- (9) Barbay K (2009) Research evidence for the use of preoperative exercise in patients preparing for total hip or total knee arthroplasty. *Orthop Nurs* 28: 127-133.
- (10) Valkenet K, van de Port IG, Dronkers JJ, de Vries WR, Lindeman E, Backx FJ (2011) The effects of preoperative exercise therapy on postoperative outcome: a systematic review. *Clin Rehabil* 25: 99-111.
- (11) Herbert RD, Bo K (2005) Analysis of quality of interventions in systematic reviews. *BMJ* 331: 507-509.
- (12) Pasquina P, Tramer MR, Walder B (2003) Prophylactic respiratory physiotherapy after cardiac surgery: systematic review. *BMJ* 327: 1379.
- (13) Lingard EA, Katz JN, Wright EA, Sledge CB (2004) Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am* 86-A: 2179-2186.
- (14) Kennedy LG, Newman JH, Ackroyd CE, Dieppe PA (2003) When should we do knee replacements? *Knee* 10: 161-166.
- (15) Santaguida PL, Hawker GA, Hudak PL, Glazier R, Mahomed NN, Kreder HJ, Coyte PC, Wright JG (2008) Patient characteristics affecting the prognosis of total hip and knee joint arthroplasty: a systematic review. *Can J Surg* 51: 428-436.
- (16) Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH, Katz JN (2002) The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol* 29: 1273-1279.

- (17) Fitzgerald JD, Orav EJ, Lee TH, Marcantonio ER, Poss R, Goldman L, Mangione CM (2004) Patient quality of life during the 12 months following joint replacement surgery. *Arthritis Rheum* 51: 100-109.
- (18) Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI, Arenaza JC, Gutierrez LF (2007) Effect of patient characteristics on reported outcomes after total knee replacement. *Rheumatology (Oxford)* 46: 112-119.
- (19) Zeni JA, Jr., Snyder-Mackler L (2010) Preoperative predictors of persistent impairments during stair ascent and descent after total knee arthroplasty. *J Bone Joint Surg Am* 92: 1130-1136.
- (20) Hulzebos EH, Helders PJ, Favie NJ, De Bie RA, Brutel dR, van Meeteren NL (2006) Preoperative intensive inspiratory muscle training to prevent postoperative pulmonary complications in high-risk patients undergoing CABG surgery: a randomized clinical trial. *JAMA* 296: 1851-1857.
- (21) Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Jr., Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS (2011) 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc* 43: 1575-1581.
- (22) Kraemer WJ, Adams K, Cafarelli E, Dudley GA, Dooly C, Feigenbaum MS, Fleck SJ, Franklin B, Fry AC, Hoffman JR, Newton RU, Potteiger J, Stone MH, Ratamess NA, Triplett-McBride T (2002) American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. *Med Sci Sports Exerc* 34: 364-380.
- (23) Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M (2008) Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 337: a1655.
- (24) Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 339: b2700.
- (25) Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 339: b2535.
- (26) Yates SL, Morley S, Eccleston C, de CWA (2005) A scale for rating the quality of psychological trials for pain. *Pain* 117: 314-325.
- (27) Robinson KA, Dickersin K (2002) Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 31: 150-153.
- (28) Higgins, J. P. T. and Green, S. (2011) *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- (29) Rohatgi A (2011) WebPlotDigitizer Available from <http://arohatgi.info/WebPlotDigitizer/>.
- (30) Boule NG, Haddad E, Kenny GP, Wells GA, Sigal RJ (2001) Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: a meta-analysis of controlled clinical trials. *JAMA* 286: 1218-1227.
- (31) van Rijn RM, van OJ, Luijsterburg PA, van MM, Koes BW, Bierma-Zeinstra SM (2010) Effectiveness of additional supervised exercises compared with conventional treatment alone in patients with acute lateral ankle sprains: systematic review. *BMJ* 341: c5688.

- (32) van Tulder MW, Suttorp M, Morton S, Bouter LM, Shekelle P (2009) Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine (Phila Pa 1976)* 34: 1685-1692.
- (33) Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174.
- (34) Chinn S (2000) A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med* 19: 3127-3131.
- (35) Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, Williams JW, Jr., Meerpohl J, Norris SL, Akl EA, Schunemann HJ (2011) GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol* 64: 1277-1282.
- (36) Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR (2000) Empirical assessment of effect of publication bias on meta-analyses. *BMJ* 320: 1574-1577.
- (37) Beaupre LA, Lier D, Davies DM, Johnston DB (2004) The effect of a preoperative exercise and education program on functional recovery, health related quality of life, and health service utilization following primary total knee arthroplasty. *J Rheumatol* 31: 1166-1173.
- (38) D'Lima DD, Colwell CW, Jr., Morris BA, Hardwick ME, Kozin F (1996) The effect of preoperative exercise on total knee replacement outcomes. *Clin Orthop Relat Res* 174-182.
- (39) Evgeniadis G, Beneka A, Malliou P, Mavromoustakos S, Godolias G (2008) Effects of pre- and postoperative therapeutic exercise on the quality of life, before and after total knee arthroplasty for osteoarthritis. *J Back Musculoskelet* 21: 161-169.
- (40) Ferrara PE, Rabini A, Maggi L, Piazzini DB, Logroscino G, Magliocchetti G, Amabile E, Tancredi G, Aulisa AG, Padua L, Aprile I, Bertolini C (2008) Effect of pre-operative physiotherapy in patients with end-stage osteoarthritis undergoing hip arthroplasty. *Clin Rehabil* 22: 977-986.
- (41) Gilbey HJ, Ackland TR, Wang AW, Morton AR, Troughet T, Tapper J (2003) Exercise improves early functional recovery after total hip arthroplasty. *Clin Orthop Relat Res* 193-200.
- (42) Gocen Z, Sen A, Unver B, Karatosun V, Gunal I (2004) The effect of preoperative physiotherapy and education on the outcome of total hip replacement: a prospective randomized controlled trial. *Clin Rehabil* 18: 353-358.
- (43) Hoogbeem TJ, Dronkers JJ, van den Ende CH, Oosting E, van Meeteren NL (2010) Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized pilot trial. *Clin Rehabil* 24: 901-910.
- (44) Rodgers JA, Garvin KL, Walker CW, Morford D, Urban J, Bedard J (1998) Preoperative physical therapy in primary total knee arthroplasty. *J Arthroplasty* 13: 414-421.
- (45) Rooks DS, Huang J, Bierbaum BE, Bolus SA, Rubano J, Connolly CE, Alpert S, Iversen MD, Katz JN (2006) Effect of preoperative exercise on measures of functional status in men and women undergoing total hip and knee arthroplasty. *Arthritis Rheum* 55: 700-708.
- (46) Topp R, Swank AM, Quesada PM, Nyland J, Malkani A (2009) The effect of prehabilitation exercise on strength and functioning after total knee arthroplasty. *PM R* 1: 729-735.

- (47) Weidenhielm L, Mattsson E, Brostrom LA, Wersall-Robertsson E (1993) Effect of preoperative physiotherapy in unicompartmental prosthetic knee replacement. *Scand J Rehabil Med* 25: 33-39.
- (48) Williamson L, Wyatt MR, Yein K, Melton JT (2007) Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement. *Rheumatology (Oxford)* 46: 1445-1449.
- (49) Wright AA, Cook CE, Flynn TW, Baxter GD, Abbott JH (2011) Predictors of response to physical therapy intervention in patients with primary hip osteoarthritis. *Phys Ther* 91: 510-524.
- (50) Veenhof C, van den Ende CH, Dekker J, Kiiike AJ, Oostendorp RA, Bijlsma JW (2007) Which patients with osteoarthritis of hip and/or knee benefit most from behavioral graded activity? *Int J Behav Med* 14: 86-91.
- (51) Hoeksma HL, Dekker J, Runday HK, Breedveld FC, van den Ende CH (2005) Manual therapy in osteoarthritis of the hip: outcome in subgroups of patients. *Rheumatology (Oxford)* 44: 461-464.
- (52) McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS (2000) Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA* 284: 79-84.
- (53) Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P (2008) Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 148: 295-309.
- (54) Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gotzsche PC, Lang T (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 134: 663-694.
- (55) Schmidt H, Mehring S, McMillan J (2010) Interpreting the declaration of Helsinki (2008): "must", "should" and different kinds of obligation. *Med Law* 29: 565-591.
- (56) Liu CK, Fielding RA (2011) Exercise as an intervention for frailty. *Clin Geriatr Med* 27: 101-110.
- (57) Glasziou P, Irwig L, Mant D (2005) Monitoring in chronic disease: a rational approach. *BMJ* 330: 644-648.
- (58) Kettunen JA, Kujala UM (2004) Exercise therapy for people with rheumatoid arthritis and osteoarthritis. *Scand J Med Sci Sports* 14: 138-142.
- (59) Chodzko-Zajko WJ, Proctor DN, Fiatarone Singh MA, Minson CT, Nigg CR, Salem GJ, Skinner JS (2009) American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Med Sci Sports Exerc* 41: 1510-1530.
- (60) Hollis S, Campbell F (1999) What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 319: 670-674.
- (61) Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 17: 1-12.
- (62) Umpierre D, Ribeiro PA, Kramer CK, Leitao CB, Zucatti AT, Azevedo MJ, Gross JL, Ribeiro JP, Schaan BD (2011) Physical activity advice only or structured exercise training

- and association with HbA1c levels in type 2 diabetes: a systematic review and meta-analysis. *JAMA* 305: 1790-1799.
- (63) Dauty M, Genty M, Ribinik P (2007) Physical training in rehabilitation programs before and after total hip and knee arthroplasty. *Ann Readapt Med Phys* 50: 462-61.
 - (64) Jack S, West M, Grocott MP (2011) Perioperative exercise training in elderly subjects. *Best Pract Res Clin Anaesthesiol* 25: 461-472.
 - (65) Wallis JA, Taylor NF (2011) Pre-operative interventions (non-surgical and non-pharmacological) for patients with hip or knee osteoarthritis awaiting joint replacement surgery - a systematic review and meta-analysis. *Osteoarthritis Cartilage* 19: 1381-1395.
 - (66) Daanen HA, Lamberts RP, Kallen VL, Jin A, van Meeteren NL (2011) A systematic review on heart rate recovery to monitor changes in training status in athletes. *IJSPP* Accepted.

11

Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized clinical pilot trial.

ABSTRACT

Objective. To evaluate the feasibility and preliminary effectiveness of therapeutic exercise before total hip replacement in frail elderly.

Design. A single-blind, randomized clinical pilot trial.

Setting. Outpatient physiotherapy department.

Subjects. Frail elderly with hip osteoarthritis awaiting total hip replacement.

Interventions. A short (3-6 weeks) tailor-made, therapeutic exercise program was compared to usual care.

Main measures. Feasibility was assessed through patient satisfaction, adherence, occurrence of adverse events and the number of eligible non-volunteers. Preliminary preoperative effectiveness was assessed with performance and self-reported measures of pain, functioning, physical activity and quality of life. Postoperative we measured functional recovery and length of hospital stay.

Results. Sixty-two eligible patients were approached of whom 21 (mean age 76 years) agreed to participate. Exercisers (n=10), rated the intervention as very well 8.9 (8-10) (10-point rating scale). No serious adverse events occurred. Forty-one (66%) eligible patients did not participate, mainly due to logistic considerations, resulting in selection bias. The intervention group (trainings sessions participated 91%) showed relevant preoperative improvements on the chair-rise time (Delta: -2.9 seconds; 95%-Confidence Interval, -6.2 to 0.4) and timed-up and go (Delta: -4.4 seconds; 95%-Confidence Interval, -9.3 to 0.5). Postoperative, no significant differences were seen.

Conclusions A short, tailor-made, exercise program is well tolerated and appreciated in elderly patients awaiting total hip replacement. However, a larger randomized clinical trial in the same setting is not warranted, due to the high number of eligible non-volunteers.

The trial registration number: ISRCTN09005050.

Hoogeboom TJ, Dronkers JJ, van den Ende CHM, Oosting E, van Meeteren NLU.

Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized clinical pilot trial.

Clinical Rehabilitation 2010;24(10):901-910.

INTRODUCTION

The number of primary total hip replacement per year is huge and is expected to grow over the next years [1]. Around 25% of the patients on the waiting list for elective hip replacement experience further loss of functional health status [2]. Knowing this and considering the growing body of literature [3-6] that suggests that decreased functional health status negatively affects the patient's postoperative functional status, it is important to prevent such decline. The hypothesis of this study is that the waiting-list period can be used for prehabilitative interventions by physiotherapists, as short term therapeutic exercise interventions have been proven to be effective in an elderly population, on condition that exercise intensity and volume are sufficient [7;8].

This hypothesis has already been studied extensively [9-14]. However the effects of a preoperative training programme on the outcome of hip replacement cannot adequately be determined [15;16], because of the limited therapeutic relevance of the studies. For example, some studies specifically excluded frail elderly patients [11;13], and some lacked the progression necessary overload during the exercise sessions [11;13]. One trial on high-risk/frail patients is promising, because their trained clients achieved discharge criteria earlier and had a shorter actual length of stay [9] but these were significantly younger than the people in the control group.

Considering the latter, we decided to develop a study according to therapeutic and methodological recommendations for quality [17]. Therapeutic quality was warranted by the specific inclusion of frail elderly patients and exclusion of relatively fit patients, sufficient exercise overload, use of intermediate outcomes [18] and evaluation on both patient perceived and performance based clinically relevant functional parameters [19]. However, before embarking on a large randomized controlled trial, we performed pilot to test the feasibility of the approach and to help clarify decisions about operational definitions and procedures [17;20]. The aims of this pilot study are to evaluate the feasibility and preliminary effectiveness of a therapeutic intervention during the waiting list period for elective total hip replacement on preoperative functioning and postoperative recovery, in frail elderly patients.

METHODS

This study was a single-blind randomized controlled pilot trial to evaluate the feasibility and preliminary effectiveness of a preoperative therapeutic exercise program compared to usual care for frail patients awaiting hip replacement. The study was approved by the

Regional Review Board situated in the University Medical Centre Utrecht and the Institutional Review Board Gelderse Vallei Hospital (protocol number 07/051).

Twenty-one frail elderly patients were recruited between July 2007 and November 2008 from the orthopedic department of the Gelderse Vallei hospital in Ede. Patients (≥ 70 years) scheduled for elective primary total hip replacement, due to end-stadium osteoarthritis, with a minimum waiting time of three weeks and ≥ 2 on the Clinical Frailty Scale [21], were eligible for inclusion. Patients that were unable to communicate in Dutch or had severe heart diseases [22] were excluded. Possible eligible patients (older 70 and osteoarthritis of the hip) were prescreened by an anesthetist for severe heart diseases. The orthopedic nurse then screened all eligible patients on our in- and exclusion criteria and informed the eligible patients about the study (both verbally and by letter). In case the patient wanted to participate in the study, the orthopedic nurse arranged a meeting with the physiotherapist. During the first session participants' questions were answered, participants gave written informed consent and participants completed the baseline assessment. Subsequently, participants were randomly allocated by use of a sealed envelope method by an independent person to either the intervention or usual care group, stratified for gender.

Both groups received preoperative and postoperative care as usual, however the intervention group also received a preoperative therapeutic intervention. The first author and two experienced physiotherapists - blinded for the patients' group allocation and trained in the measurement techniques - administered the baseline and preoperative measures, respectively. The measurements during the hospital stay were also blindly administered by experienced and trained physiotherapists. The therapeutic intervention was provided by three other physiotherapists.

In the intervention group participants trained at least two times per week supervised at the outpatient department of physiotherapy for a minimum of three and a maximum of six weeks before surgery. Adjacent to the supervised exercise, patients were encouraged to exercise at home. Each supervised training session consisted of 60 minutes and was composed of four phases. First, patients started with a five minute walk to warm-up. Subsequently, they trained their lower extremity with the leg-press (sets of 10-20) through the full possible range of motion, both concentric as eccentric [23]. Then, patients trained their aerobic capacity on a bicycle ergo meter for 20-30 minutes. Finally, they followed a specific tailor-made training which integrates functional physical activities into the patient's daily living [24]. The exercise intensity was set to 13-14 (moderate to high) on a 15-point rated perceived exertion scale (6 = very, very light; 20 = very, very hard) [25-27], except for the warm-up. If an exercise was rated 'somewhat hard' (12 on rated perceived exertion-scale), participants were instructed to increase the weight carried, the number of repetitions, and/or the distance walked [22;27;28]. If necessary, manipulative tractions were used to reduce any discomfort the participants were experiencing [29]. On non-training days the participants were stimulated to expand their normal activity pattern [30], by use of a pedometer. To facilitate adherence to the program we monitored progression [18], let

patients include their spouse/family in the exercises, implemented feedback moments about the intensity of and attitude towards the exercise, and progressed the training parameters gradually.

Usual care in the preoperative phase consisted of one group-based education session about early mobilization, surgery and anesthesia techniques, restricted movements, benefits of activity, and proper use of crutches approximately one week before surgery. The postoperative usual care protocol is strict. Operation day (day 0): the patient exercises, if possible, in bed. Day after the operation (day 1): The patient gets out of bed with their physiotherapist and an orthopedic nurse. Day 2 till discharge: The patient starts walking with a walking-aid (usually two crutches) and climbing stairs. These objectives are normally met after 5 to 7 days, unless a patient rehabilitates in a nursery home (discharge at the forth day) or has serious complications.

Data on feasibility and preliminary effectiveness were collected within three days before the start of the exercise program (baseline), within a week of surgery (preoperative) and during each day of the hospital stay (postoperative) by a blinded assessor. At baseline, next to demographic data, two additional frailty measures were assessed, namely maximal hand grip strength [31] and two tests developed by Gill *et al* (1995) [32], to determine the validity of our frailty cut-off score.

Our primary outcome measure was feasibility of the intervention. Feasibility measures were occurrence of adverse events during testing or training, the number of dropouts, adherence to exercise, and patients' satisfaction with the treatment determined. To investigate the feasibility of the study, we compared the demographic data from both the volunteers and the eligible non-volunteers. Significant differences between these two groups would indicate that the study was not feasible.

Secondary outcome measures to evaluate the preoperative effectiveness of the intervention, were both performance based and self-reported. The Hip Osteoarthritis Outcome Score [33], a self-reported questionnaire with good psychometric properties was utilized to measure pain, stiffness, physical functioning, functioning in sport and recreation and hip related quality of life [34]. For each scale a score of 100 indicates no symptoms and 0 indicates extreme symptoms. Self-reported physical activity was measured with the LASA Physical Activity Questionnaire [35]. The LASA Physical Activity Questionnaire measures the frequency and the number of minutes a day that are spend on activities like cycling, gardening and walking, and is both reliable and valid for classifying physical activity in an older population [35]. Physical functioning assessed with the patient specific complaints questionnaire. This patient-specific questionnaire asks patients to select three activities that (s)he perceives as problematic (activities that can easily be avoided are not allowed) and scores the severity on a 10 cm visual analogue scale [36]. Physical work capacity was measured with the PWC-170 ergometer test [37]. During this test subjects are asked to cycle at a constant 60 rpm for three minutes at three successively increasing but sub-maximal workloads. Heart rate and rated perceived exertion were recorded at one and three

minute intervals at each workload, respectively. Physical work capacity at 170 beats/minute will be estimated by extrapolating the line of best fit to the three sub-maximal heart-rate-work-capacity data points to estimate PWC-170. Walking capacity was measured with the six minute walk test [19]. For the six minute walk test the patient is asked to cover as much distance walking within 6 minutes, but is allowed to take a break when needed. It was executed over a trajectory of 30 meters. Intraclass correlation for test-retest reliability is good; 0.94 (0.88, 0.98) [19]. Functional mobility was measured with the timed up and go test [38]. During the timed up and go test the patient is observed and timed while he rises from an arm chair, walks three meters, turns, walks back, and sits down again. This test is found reliable and valid for quantifying functional mobility in frail elderly [38]. Functional extension power of the legs was measured with the chair rise time. For the chair rise time score the time is measured that patients need to rise from a chair five times with the arms folded across the chest. Use of the chair arms was permitted if unable to stand with arms positioned across the chest.

Postoperative effectiveness was assessed by length of stay and postoperative recovery. Postoperative recovery was evaluated with the reliable, valid and responsive [39;41] Iowa Level of Assistance Scale [39;40]. The scale rates specific tasks of transfer, ambulation and step negotiation from independent (0), through supervision (1), to assistance (2-4), to unable (5) and unsafe to test (6). A total score of one or nil means that the patient is able to function in their daily living without any assistance; if a patient does not need or does not want to climb stairs, they reach the theoretical discharge at a score of six.

For the statistical analysis, we stored all data in Excel (Microsoft® Office, 2003) and analyzed the data with SPSS (version 15.0) statistical software (SPSS Inc., Chicago, IL, USA). The data were then checked for completeness and normality. The analysis was performed according to the intention-to-treat principle. Summary descriptive statistics, including frequencies, means, standard deviations and percentages were computed for the preoperative variables. To evaluate the effectiveness of the therapy, all outcome measures were analyzed using analysis of covariance, on the condition that all assumptions are fulfilled. In this analysis, we will adjust the preoperative scores for the baseline scores, according to recommendations of Vickers *et al* [42]. Postoperative functional recovery and length of stay were described with descriptive statistics and clustered histograms. Treatment effects are represented by point estimates and 95% confidence intervals. The analysis of the serial measures of functional recovery was performed with a two stage method that uses summary measures [43]. Postoperative functional recovery decrease steadily with time and describes a 'growth curve'. To determine any delays in postoperative functional recovery between the two groups, we performed a Mann-Whitney U test on the time the participants needed to reach their functional discharge criteria (determined with the Iowa Level of Assistance score).

RESULTS

During the inclusion period 62 patients met all eligibility criteria (Figure 1). Twenty-one patients were willing and able to participate, while the other 41 patients declined (main reasons: no transport (n=17), no energy (n=10), no time (n=6), not motivated (n=4) and some already had therapy (n=4)). The eligible non-volunteers were significantly ($p=0.002$) older (mean \pm sd age: 80 \pm 5 years) than the participants in the study (76 \pm 4 years).

Table 1. Baseline characteristics of study participants.

	Study group (N=10)	Usual care group (N=11)
Female/Male	7/3	7/4
Age at surgery, years (mean \pm SD, range)	77 \pm 3 (71-82)	75 \pm 5 (69-90)
Days on waiting list (mean \pm SD, range)	29 \pm 7 (18-44)	34 \pm 14 (14-63)
HOOS function, daily activities (mean \pm SD)	51.5 \pm 16.7	54.8 \pm 19.7
BMI, kg/m ² (mean \pm SD, range)	26.0 \pm 2.6 (22.2-30.1)	27.4 \pm 4.2 (25.9-35.9)
Number of CMs (median, range)	1.5 (0-3)	1 (0-4)
Approach, posterolateral / anterior	6/4	8/3
Bilateral complaints, yes/no	3/7	0/11
Preoperative help, yes/no	5/5	8/3
Postoperative help, yes/no	7/3	4/7
CFS-score (median, range)	2 (2-5)	3 (2-5)

Abbreviations: BMI=Body Mass Index, CMs=Comorbidities, CFS=Clinical Frailty Scale, HOOS = Hip Osteoarthritis Outcome Score, SD = Standard Deviation.

No differences were found between the exercise group (n=10) and the usual care group (n=11) at baseline (table 1). The average baseline characteristics of both groups are listed in Table 1.

Patients in the exercise group participated 91% of all sessions. During the intervention, no serious adverse events occurred. All ten participants valued the therapy, the received information and the treatment as very well (mean overall score (range) 8.9 (7-10), 8.3 (7-10) and 9.3 (7-10), respectively) (10-point scale). The overall contentment with the physiotherapists and the outpatient clinic was 8.9 (8-10). Furthermore, participants estimated the effect of the treatment before they the exercise on a 7.5 (6-9) and after the training on a 8.9 (8-10), $p=0.001$.

The baseline measures and the differences between preoperative means in self-rated measures and in performance measures are listed in Table 2 and Table 3, respectively. No significant differences between groups in any of the self-reported measures were revealed.

Figure 1. Flow chart patient recruitment. The two-sided arrows represent comparative statistical analyses.

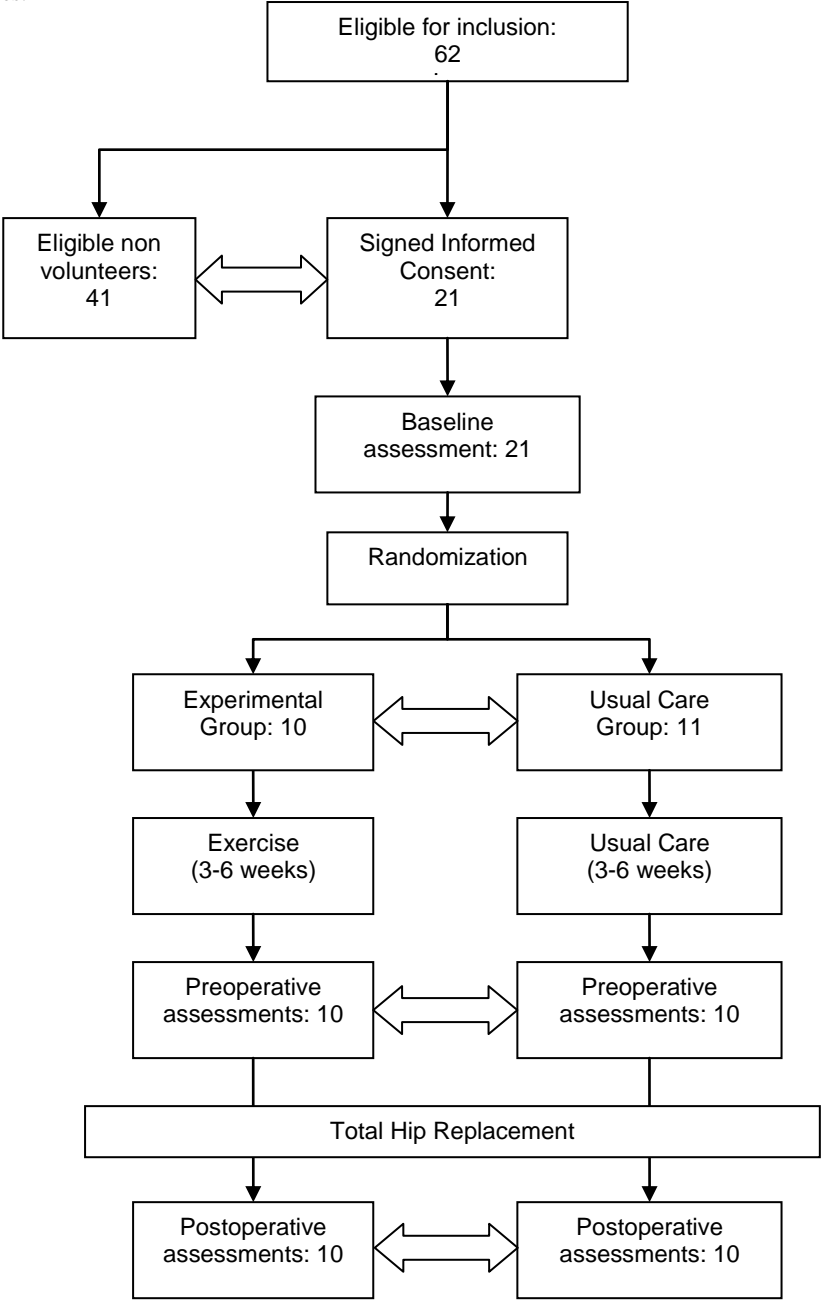


Table 2. Scores for the self-rated outcome measures of all participants.

	Exercise Mean ± SD	Usual Care Mean ± SD	Between-group differences*
HOOS Pain			
Baseline	48.0 ± 16.7	53.0 ± 17.7	7.3 (-5.1 to 19.7)
Preoperative	55.3 ± 12.0	49.3 ± 17.0	
HOOS Symptoms			
Baseline	57.5 ± 13.0	59.5 ± 14.6	-0.6 (-12.2 to 11.0)
Preoperative	56.5 ± 14.2	59.0 ± 15.6	
HOOS Function, daily activities			
Baseline	51.5 ± 16.7	54.8 ± 19.7	-0.6 (-11.9 to 10.6)
Preoperative	51.1 ± 10.5	52.3 ± 21.2	
HOOS Function, sport & recreation			
Baseline	22.5 ± 15.6	27.8 ± 20.8	-4.6 (-14.0 to 4.9)
Preoperative	25.0 ± 14.4	32.5 ± 20.2	
HOOS Quality of live			
Baseline	30.7 ± 13.6	41.0 ± 18.3	-1.2 (-13.0 to 10.7)
Preoperative	36.3 ± 15.8	43.3 ± 15.4	
VAS Pain			
Baseline	36.0 ± 20.8	45.8 ± 26.8	-8.6 (-27.3 to 10.2)
Preoperative	37.7 ± 18.0	49.3 ± 24.7	
PSK (total score of 5 VAS scales)			
Baseline	316.3 ± 85.7	273.9 ± 66.7	1.9 (-54.1 to 57.9)
Preoperative	306.9 ± 88.3	265.2 ± 103.8	
LAPAQ (in Kcal/day)			
Baseline	671.6 ± 345.5	615.1 ± 397.6	69.2 (-207.1 to 345.5)
Preoperative	706.6 ± 302.8	631.5 ± 365.6	
LAPAQ (in Minutes)			
Baseline	203.3 ± 117.1	186.1 ± 116.6	-4.2 (-87.7 to 79.3)
Preoperative	213.7 ± 81.9	216.4 ± 134.6	
LAPAQ (in Kcal/Minutes/Weight)			
Baseline	0.05 ± 0.01	0.04 ± 0.01	0.0 (-0.1 to 0.1)
Preoperative	0.04 ± 0.01	0.04 ± 0.01	

*Analysis of covariance (adjusted mean difference (95% CI)). Abbreviations: HOOS = Hip disability and Osteoarthritis Outcome Score (0=worst, 100=best), LAPAQ = LASA Physical Activity Questionnaire, Kcal = kilocalorie, PSK = Patient Specific Complaints, VAS = Visual Analogue Scale (0=best, 100=worst).

However, relevant differences between groups were found for the chair rise time test (Delta: -2.9 seconds; 95%-Confidence Interval, -6.2 to 0.4; $p=0.084$) and the timed up and

go test (Delta: -4.4 seconds; 95%-Confidence Interval, -9.3 to 0.5; $p=0.073$). Other performance measures showed no significant differences.

No differences were found in the postoperative length of stay. The median (range) length of stay was 6 days (5-22) for the intervention ($n=10$) and 6 days (4-7) for the usual care group ($n=10$) ($p=0.228$). Also, no differences were found in the functional recovery - described as the time needed to reach functional independence - between the intervention group and the usual care group ($p=0.963$); 4 days (3-6) vs 4 days (3-5). Note that not all patients reached functional independence, due to an early transfer to another institute ($n=1$) and postoperative complications ($n=2$).

Postoperative complications occurred in two participants (both exercisers). One participant fractured her femur during surgery and was instructed to stay in bed for at least 6 weeks. Another participant had a low saturation rate after surgery (89%), due to this the patient was restricted from walking; on the fourth day he was diagnosed with intestinal obstruction.

Table 3. Scores for all performance measures of all participants.

	Exercise mean \pm SD	Usual Care mean \pm SD	Between-group differences*
6MWT (in meters)			
Baseline	359,7 \pm 117.4	336.8 \pm 92.1	5.3 (-43.7 to 54.3)
Preoperative	363.0 \pm 126,3	342.7 \pm 133.7	
TUG (in seconds)			
Baseline	11.3 \pm 4.7	10.9 \pm 4.5	-4.4 (-9.3 to 0.5)
Preoperative	9.6 \pm 2.8	13.4 \pm 11.5	
CRT (in seconds)			
Baseline	18.5 \pm 13.5	17.1 \pm 6.4	-2.9 (-6.2 to 0.4)
Preoperative	15.0 \pm 5.8	17.4 \pm 5.9	
PWC-170			
Baseline	99.6 \pm 146.1	86.22 \pm 45.0	-17.9 (-92.6 to 56.9)
Preoperative	103.6 \pm 74.5	120.7 \pm 63.3	
HGS Peak (in Newtons)			
Baseline	275.5 \pm 107.2	306.0 \pm 113.0	-12.7 (-47.3 to 21.9)
Preoperative	289.6 \pm 105.8	317.9 \pm 85.1	
HGS Average (2 sec) (in Newton)			
Baseline	241.2 \pm 94.4	254.7 \pm 121.7	-35.7 (-88.2 to 16.7)
Preoperative	243.7 \pm 92.2	283.5 \pm 83.3	

*Analysis of covariance (adjusted mean difference (95% CI)). Abbreviations: 6MWT = Six minute walk test, TUG = Timed Up and Go test, CRT = Chair Rise Time, PWC-170 = Aerobic Bicycle Ergometer Test, HGS = Hand Grip Strength.

DISCUSSION

In this pilot study we found that the current study design and procedures do not justify a larger randomized controlled trial, as we only recruited 34% of the eligible patients. On the other hand, the intervention appeared feasible as none of the exercisers dropped out, all participants were very satisfied, adherence rate was high and no serious adverse events occurred other than those that expected [44]. Considering the preliminary effectiveness of the trial, outcomes were promising, even though no significant differences were seen on postoperative outcome measures between the two groups in this pilot study, the intervention group did show relevant preoperative improvement on two relevant functional measures; chair rise time and timed up and go test.

Feasibility of the study design was low, due to the slow and incomplete recruitment of elderly patients to the study. Only 34% of the eligible patients were recruited, as most patients were unable to drive a car and were dependent of friends or spouse to take them to the hospital twice a week. These findings are similar to previous studies [12;14;32]. Because of the low inclusion rate, selection bias occurred; eligible non-volunteers were significantly older than volunteers in the study. To prevent this form of selection bias in future research, we believe that therapists should provide the preoperative therapeutic exercise at home. This, however, must again be tested in a pilot study. Feasibility should – besides all other aspects of a pilot study – in this particular case concentrate on avoiding such type of selection bias. On the other hand, we found the exercise intervention as such to be feasible for those that did participate. Previous studies on the matter reported good acceptance of the therapy, high attendance rates, high degree of enjoyment and perceived value of the intervention [15;16]; findings similar to our results.

A limitation in our study was the frailty screening. As far as we know, no clinical prediction rules were available to assess which patients are predisposed for delayed functional recovery after total hip replacement. In this study the clinical frailty scale was used to assess frailty, as this is an easy to use scale, which correlates good with other frailty scales [21]. At the start of the study, the clinical frailty scale was set at ≥ 5 (5=mildly frail), though after 10 weeks 3 participants were included, the scale cut-off was set back to ≥ 2 . The reason for this alteration was that we believed this cut-off point was not sensitive for patients with physical impairments as the clinical frailty scale mainly differentiates on comorbid diseases instead of the patient's physical capacity. Though in retrospect, this alteration of eligibility criteria let more physically fit people enroll in the study. For example, the timed up and go times differed 4 seconds and the six minute walk test 70 meters in favor of the patients included after these 10 weeks. According to the criteria set by Gill *et al* (1995) [32] only five participants could be called physically frail (three exercisers and two usual care patients) and none of the participants were frail according to

the hand grip strength [31], resulting in a total of seven patients (33%) with objectified frailty in at least one of the three definitions (Table 4).

Table 4. Individual baseline scores of patients included in the study.

	Number of patients frail according to CFS ≥ 5 [21]	Number of patients frail according to Gill [32]	Number of patients frail according to HGS [31]	Total number of frail patients according to CFS, Gill or HGS*
Exercise group	3/10	2/10	0/10	3/10
Usual care group	2/11	2/11	0/11	4/11
Total	5	4	0	7/21

*Note that there is overlap in patients between frailty definitions. Abbreviations: CFS = Clinical Frailty Scale, HGS = Hand Grip Strength.

Thus, considering the wide variation of patients included in this study, it remains to be seen whether the persons selected on the base of their clinical frailty scale-score indeed were at risk for postoperative delayed recovery. The latter emphasizes the need for proper selection criteria to detect the frail patients and for a well-designed study that assess the relation between patients' preoperative health status and their postoperative recovery and ends up in a clinical prediction rule that can be used and tested for its validity [45].

The methodological strength of this pilot study is its randomized, single blind, controlled design. Another strength of this study is its clinical relevance. Two systematic reviews that investigated preoperative physical therapy [15;16] and multidisciplinary rehabilitation [16], found undetermined benefits for total hip replacement on postoperative length of stay or discharge criteria, but emphasized the need for more clinically relevant studies. This pilot study [a] aimed to include patients at risk for delayed recovery, [b] provided a tailor-made therapeutic exercise intervention, [c] induced sufficient overloading during the exercise session by titration of the intensities of the exercises in order to optimize the functional and conditional gain (as is normally done in a therapeutic setting), [d] monitored the immediate success of intervention on functional and conditional progression with intermediate outcomes, and [e] assessed, beside the length of stay, also the postoperative functional recovery.

In conclusion, we believe the results obtained in this pilot trial do not justify the execution of a larger randomized clinical trial and a secondary pilot trial is needed to evaluate the effect of a home-based exercise program on the inclusion rate of patients. On the other hand, our data show that elderly patients can be trained without serious side effects and/or adverse events, but one has to stick to principles of real therapeutic practice. So, future studies must consider the high number of eligible non-volunteers and adjust their study logistics / design / intervention accordingly.

REFERENCE LIST

- (1) Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am* 2007 Apr;89(4):780-5.
- (2) Hoogboom TJ, van den Ende CH, van der Sluis G, et al. The impact of waiting for total joint replacement on pain and functional status: a systematic review. *Osteoarthritis Cartilage* 2009 May 20.
- (3) Fortin PR, Clarke AE, Joseph L, et al. Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery. *Arthritis Rheum* 1999 Aug;42(8):1722-8.
- (4) Fortin PR, Penrod JR, Clarke AE, et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. *Arthritis Rheum* 2002 Dec;46(12):3327-30.
- (5) Hajat S, Fitzpatrick R, Morris R, et al. Does waiting for total hip replacement matter? Prospective cohort study. *J Health Serv Res Policy* 2002 Jan;7(1):19-25.
- (6) Holtzman J, Saleh K, Kane R. Effect of baseline functional status and pain on outcomes of total hip arthroplasty. *J Bone Joint Surg Am* 2002 Nov;84-A(11):1942-8.
- (7) Gibala MJ, McGee SL. Metabolic adaptations to short-term high-intensity interval training: a little pain for a lot of gain? *Exerc Sport Sci Rev* 2008 Apr;36(2):58-63.
- (8) Wisloff U, Stoylen A, Loennechen JP, et al. Superior cardiovascular effect of aerobic interval training versus moderate continuous training in heart failure patients: a randomized study. *Circulation* 2007 Jun 19;115(24):3086-94.
- (9) Crowe J, Henderson J. Pre-arthroplasty rehabilitation is effective in reducing hospital stay. *Can J Occup Ther* 2003 Apr;70(2):88-96.
- (10) Gilbey HJ, Ackland TR, Wang AW, Morton AR, Trouchet T, Tapper J. Exercise improves early functional recovery after total hip arthroplasty. *Clin Orthop Relat Res* 2003 Mar;408:193-200.
- (11) Gocen Z, Sen A, Unver B, Karatosun V, Gunal I. The effect of preoperative physiotherapy and education on the outcome of total hip replacement: a prospective randomized controlled trial. *Clin Rehabil* 2004 Jun;18(4):353-8.
- (12) Rooks DS, Huang J, Bierbaum BE, et al. Effect of preoperative exercise on measures of functional status in men and women undergoing total hip and knee arthroplasty. *Arthritis Rheum* 2006 Oct 15;55(5):700-8.
- (13) Wijnman AJ, Dekkers GH, Waltje E, Krekels T, Arens HJ. [No positive effect of preoperative exercise therapy and teaching in patients to be subjected to hip arthroplasty]. *Ned Tijdschr Geneesk* 1994 May 7;138(19):949-52.
- (14) Ferrara PE, Rabini A, Maggi L, et al. Effect of pre-operative physiotherapy in patients with end-stage osteoarthritis undergoing hip arthroplasty. *Clin Rehabil* 2008 Oct;22(10-11):977-86.
- (15) Ackerman IN, Bennell KL. Does pre-operative physiotherapy improve outcomes from lower limb joint replacement surgery? A systematic review. *Aust J Physiother* 2004;50(1):25-30.

- (16) Coudeyre E, Jardin C, Givron P, Ribinik P, Revel M, Rannou F. Could preoperative rehabilitation modify postoperative outcomes after total hip and knee arthroplasty? Elaboration of French clinical practice guidelines. *Ann Readapt Med Phys* 2007 Apr;50(3):189-97.
- (17) Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655.
- (18) Glasziou P, Irwig L, Mant D. Monitoring in chronic disease: a rational approach. *BMJ* 2005 Mar 19;330(7492):644-8.
- (19) Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskelet Disord* 2005;6:3.
- (20) Portney L, Watkins M. The research question. *Foundations of clinical research*. Prenti-Hall; 2000.
- (21) Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005 Aug 30;173(5):489-95.
- (22) American College of Sports Medicine Position Stand. Exercise and physical activity for older adults. *Med Sci Sports Exerc* 1998 Jun;30(6):992-1008.
- (23) Mazzeo RS, Tanaka H. Exercise prescription for the elderly: current recommendations. *Sports Med* 2001;31(11):809-18.
- (24) Sevick MA, Dunn AL, Morrow MS, Marcus BH, Chen GJ, Blair SN. Cost-effectiveness of lifestyle and structured exercise interventions in sedentary adults: results of project ACTIVE. *Am J Prev Med* 2000 Jul;19(1):1-8.
- (25) de Vreede PL, Samson MM, van Meeteren NL, Duursma SA, Verhaar HJ. Functional-task exercise versus resistance strength exercise to improve daily function in older women: a randomized, controlled trial. *J Am Geriatr Soc* 2005 Jan;53(1):2-10.
- (26) de Vreede PL, Samson MM, van Meeteren NL, van der Bom JG, Duursma SA, Verhaar HJ. Functional tasks exercise versus resistance exercise to improve daily function in older women: a feasibility study. *Arch Phys Med Rehabil* 2004 Dec;85(12):1952-61.
- (27) Robertson RJ, Goss FL, Rutkowski J, et al. Concurrent validation of the OMNI perceived exertion scale for resistance exercise. *Med Sci Sports Exerc* 2003 Feb;35(2):333-41.
- (28) Lagally KM, Robertson RJ, Gallagher KI, Gearhart R, Goss FL. Ratings of perceived exertion during low- and high-intensity resistance exercise by young adults. *Percept Mot Skills* 2002 Jun;94(3 Pt 1):723-31.
- (29) Hoeksma HL, Dekker J, Runday HK, et al. Comparison of manual therapy and exercise therapy in osteoarthritis of the hip: a randomized clinical trial. *Arthritis Rheum* 2004 Oct 15;51(5):722-9.
- (30) Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. *J Am Geriatr Soc* 2003 Mar;51(3):387-92.
- (31) Kerr A, Syddall HE, Cooper C, Turner GF, Briggs RS, Sayer AA. Does admission grip strength predict length of stay in hospitalised older patients? *Age Ageing* 2006 Jan;35(1):82-4.

- (32) Gill TM, Williams CS, Tinetti ME. Assessing risk for the onset of functional dependence among older adults: the role of physical performance. *J Am Geriatr Soc* 1995 Jun;43(6):603-9.
- (33) Klassbo M, Larsson E, Mannevik E. Hip disability and osteoarthritis outcome score. An extension of the Western Ontario and McMaster Universities Osteoarthritis Index. *Scand J Rheumatol* 2003;32(1):46-51.
- (34) de Groot I, Reijman M, Terwee CB, et al. Validation of the Dutch version of the Hip disability and Osteoarthritis Outcome Score. *Osteoarthritis Cartilage* 2007 Jan;15(1):104-9.
- (35) Stel VS, Smit JH, Pluijm SM, Visser M, Deeg DJ, Lips P. Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. *J Clin Epidemiol* 2004 Mar;57(3):252-8.
- (36) Beurskens AJ, de Vet HC, Koke AJ, et al. A patient-specific approach for measuring functional status in low back pain. *J Manipulative Physiol Ther* 1999 Mar;22(3):144-8.
- (37) Withers RT, Davies GJ, Crouch RG. A comparison of three W170 protocols. *Eur J Appl Physiol Occup Physiol* 1977 Sep 16;37(2):123-8.
- (38) Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991 Feb;39(2):142-8.
- (39) Shields RK, Enloe LJ, Evans RE, Smith KB, Steckel SD. Reliability, validity, and responsiveness of functional tests in patients with total joint replacement. *Phys Ther* 1995 Mar;75(3):169-76.
- (40) Jesudason C, Stiller K. Are bed exercises necessary following hip arthroplasty? *Aust J Physiother* 2002;48(2):73-81.
- (41) Oldmeadow LB, McBurney H, Robertson VJ. Hospital stay and discharge outcomes after knee arthroplasty: implications for physiotherapy practice. *Aust J Physiother* 2002;48(2):117-21.
- (42) Vickers AJ, Altman DG. Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ* 2001 Nov 10;323(7321):1123-4.
- (43) Matthews JN, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990 Jan 27;300(6719):230-5.
- (44) Forster A, Lambley R, Hardy J, et al. Rehabilitation for older people in long-term care. *Cochrane Database Syst Rev* 2009;(1):CD004294.
- (45) McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA* 2000 Jul 5;284(1):79-84.

12

General discussion

INTRODUCTION

Osteoarthritis (OA) is the most prevalent chronic joint disease (1). It is associated with a substantial burden to the individual (i.e. pain, stiffness, loss of function and reduced quality of life) and considerable costs to society (i.e. disability and loss of productivity). Despite the impact of OA on the individual and society, no disease-modifying medications (DMOADs) are yet available (2), although some candidate DMOADs are currently being tested clinically (3). There are, however, several pharmacological, non-pharmacological and surgical treatment options available that can help people to cope with OA. Unfortunately, (non-)pharmacological interventions only elicit small to moderate therapeutic effects (4-8) and surgical techniques (mainly total joint replacement) still result in unsatisfactory outcomes in a substantial group of people (9;10).

This thesis reports on research into tailoring health care programs to the needs of individuals with OA to optimize therapy outcomes. We assessed several (aspects of) health needs in OA by (a) investigating the validity of the most often used measure to determine health needs, (b) identifying subgroups of patients with different health needs, (c) studying the natural course of pain and functioning in end-stage OA, and (d) investigating health care utilization in context of the Dutch health care system. Moreover, we reported on (a) the development and preliminary evaluation of the feasibility and effectiveness of non-pharmacological interventions tailored to health needs in individuals with advanced OA. Section 12.1 discusses the main findings underlying this thesis. Methodological reflections are offered in section 12.2. In section 12.3 we discuss the implementation of our findings into clinical practice and in section 12.4 we give our recommendations for further research.

12.1 MAIN RESULTS

Assessing health needs in OA:

In current OA-related research the most assessed and studied health parameters are pain, stiffness and functional limitations. To date, the WOMAC scale is one of the most frequently used scales to assess pain and limitations in activities in OA. Its usefulness for clinical practice should however be questioned, as its ability to differentiate between pain and functioning (i.e. discriminant validity) is limited (11-15). In Chapter 2 we demonstrated that the Dutch version of the Lower Extremity Functional Scale (LEFS) was both highly valid and reliable and that the LEFS, in contrast to the KOOS and the HOOS (and thus also the similar itemed WOMAC) could adequately differentiate pain from functioning (16). Furthermore, the LEFS appeared to be more responsive to change than the function scales of the KOOS and HOOS.

Disappointing treatment effects could potentially be explained through effect modification by the heterogeneity of the disease as the magnitude of the effect might differ among subgroups of patients with different characteristics. We hypothesized that individuals reporting joint-pain comorbidity (JPC) are clinically different than patients without JPC. We studied our hypothesis in a single-centre, cross-sectional study and in a multi-centre, cohort study (Chapters 3 and 4) (17;18). Both studies yielded similar results, namely, individuals experiencing JPC report a higher clinical burden than those who do not experience JPC. Interestingly, we were unable to demonstrate relevant differences in health care consumption between the two groups.

Another group of individuals that might be at risk for compromised health status are those awaiting elective total hip or knee replacement. Since preoperative functioning is considered the strongest determinant of postoperative pain and functioning (19-24), it is important to understand whether prolonged preoperative delay is associated with a further deterioration in the preoperative status. By means of systematically reviewing the literature we assessed the natural course of pain and functioning during the waiting list period for major joint replacement (Chapter 5) (25). In addition, we aimed to identify determinants for decline in pain or functioning. The retrieved studies allowed us to draw only valid conclusions regarding ‘short’ waiting list periods (i.e. <180 days) and demonstrated that on average patients with OA do not experience deterioration in pain or self-reported functional status while waiting. However, a quarter of the patients did report clinically relevant deterioration during the waiting time; these individuals might benefit from a preoperative therapeutic exercise program to prepare them for surgery to optimize their postoperative functional recovery.

In the final study of the part on ‘health needs assessment’ of the thesis we looked at *health needs* in the context of the health care system. Analysing the two-year results of a multi-centre Dutch OA cohort study, we investigated the course of health care consumption in OA over time and its determinants (Chapter 6) (26). Ideally, individuals use health care modalities on the basis of their discomfort level (i.e. disease-related factors) rather than on the basis of predisposing (e.g. sex, age, ethnicity, etc) or enabling (e.g. insurance type, income, etc) factors. Generally, health care utilization was low during the first two years of the disease. Only the use of supplements increased over time, even though their efficacy is highly questionable (27;28). It was remarkable that predisposing and enabling factors contributed almost equally to analgesic use and health professional contact as disease-related factors.

Tailoring care in OA:

To develop a therapeutically valid and effective intervention several steps must be considered (Chapter 7) (29). By use of a case-study (in this case hand OA) we presented a systematic approach, as recommended by the literature, to develop an evidence-based

treatment program by synthesizing information from expert opinions, patient perspectives, and scientific literature.

Self-management for people with generalized OA

Following our insights in conceptualizing a complex intervention and our knowledge about the burden of experiencing pain in multiple joint groups, we developed a multi-disciplinary and multi-dimensional intervention for individuals with generalized OA (GOA) (30). People were considered having GOA when they experienced complaints in three or more joint groups and had at least two objective signs that indicate OA in at least two of those joints (30). In this thesis we investigated the feasibility and preliminary effectiveness of a 12 week, multi-disciplinary care program for individuals with GOA, by use of a concurrent, multiple baseline, single case study (Chapter 9) (31). Although no adverse events occurred and adherence rate was high, participant satisfaction about the therapy was moderate as well as the participants' estimate of effectiveness of the intervention. Moreover, we found no beneficial effects on pain or on self-efficacy. We concluded that the proposed intervention for patients with GOA was insufficiently feasible and effective; however, data obtained by this multiple-baseline study provided data that allowed us to optimize the therapy program (Chapter 8) (30).

Preoperative exercise for frail elderly awaiting total joint replacement

A subgroup of individuals - mainly those of older age (20;22) with low levels of preoperative functioning (19-24) - on the waiting list for elective primary hip or knee surgery is at risk for disappointing surgery outcome (9;10). We hypothesized that it might be of value to optimize the preoperative functional status of these individuals to enhance their postoperative functional recovery. To do so, we systematically reviewed the available literature on the effectiveness of preoperative exercise (Chapter 10) (32). To ensure that we only used data from high quality studies in our meta-analysis, we set *a priori* standards for methodological and therapeutic validity. Criteria for therapeutic validity - defined as *the potential effectiveness of a specific intervention given the potential target group of patients* - was determined using a Delphi-study generated rating scale (Chapter 10) (32). We found that none of the studied interventions fulfilled the criterion of therapeutic validity; aspects such as adequate patient selection, adequate dosing and monitoring were missing. When pooling these low quality studies, the benefit of preoperative therapeutic exercise on postoperative functional recovery could not be demonstrated. It remains to be seen whether a high quality study would result in beneficial effects. Before embarking on such a large trial, we performed a randomized clinical pilot trial. In this pilot study we tested a high intensity exercise program on individuals that were predisposed for delayed postoperative recovery (Chapter 11) (33). We found, that a short, tailor-made, exercise program is well tolerated and appreciated in elderly patients awaiting total hip replacement. However, due

to the high number of eligible non-volunteers, who were older and frailer, a large randomized clinical trial in the same setting is not warranted.

12.2 METHODOLOGICAL REFLECTIONS

In this section we discuss several issues that could have limited the validity of our results in this thesis.

Systematic reviews

In this thesis we report the results of two systematic reviews. Although systematic reviews are an easy and cost-efficient way to gain insight in a topic, one is always dependent of the quality of the available studies. We encountered several issues in the current literature that hampered us in drawing robust conclusions on our research questions.

We found a tendency in the literature on preoperative and postoperative patient evaluations to use self-report measures rather than observed measures to determine physical functioning. A solid body of evidence demonstrates that self-reported physical functioning is a different construct than observed physical functioning (34-36). Observed performance data of an individual's functional capacity is more predictive for postoperative in-hospital recovery than self-report data (37). So, for a clinician it would be more interesting to know whether a patient deteriorated in his/her observed functional status than self-reported functional status, and therefore we preferred performance-tested data over self-report data in our data-synthesis. However, only self-report data were available, limiting us in providing a valid answer to our research question. In addition, the validity of research findings based merely on self-report data could also be limited due to response shift and recall bias (38-40).

A common criticism on the scientific literature is that evaluated non-pharmacological therapies are of poor therapeutic quality (41), making it unlikely for them to elicit beneficial effects (42). In our review on the effectiveness of therapeutic exercise prior to major joint replacement, we indeed found that none of the included studies evaluated an intervention that matched consensus-based requirements for high quality therapeutic exercise (32). The lack of therapeutic validity in the included studies did not allow us to draw firm conclusions about the efficacy of preoperative exercising.

A final concern that needs to be addressed is the overall low methodological quality of available studies. A total of 27 studies were included in both reviews of which only four (15%) could be qualified as methodologically sound. Of course, insufficient reporting accounts for a part of these low quality studies and could perhaps have resulted in the incorrect labelling of high quality studies as being of low quality. However, study results of studies that cannot be reproduced should always be questioned.

Observational studies

This thesis consists of four studies with an observational design, one cross-sectional study (17) and three longitudinal studies (16;18;26). The major limitation of cross-sectional studies is that they are confined to a specific point in time. That is, they provide only a snapshot view of a sample of the population at a single point in time. Because disease-related characteristics can change over time, cross-sectional studies therefore might only reflect the situation at a single time point. Thus, the preferred method of evaluating associations between variables is a longitudinal design as it also takes changes within subjects into account.

A major validity threat in observational studies is selection bias. Selection bias occurs when a sample is collected in such a way that some members of the intended population are less likely to be included than others. In our cross-sectional study on JPC and our longitudinal study on the validity of the LEFS we had some concerns regarding the validity of our sample, as all participants were recruited at a single centre; the orthopaedic department of the Sint Maartenskliniek. As the Sint Maartenskliniek is a specialized hospital in the Netherlands, for instance expressed by the high number of second opinions provided, the generalisability of the study results should be interpreted with care. On the other hand, when comparing our sample with samples from other multi-centre studies (43) we did find similarities regarding demographics, warranting – to some extent – the generalizability of our findings.

Another issue that can impact the validity of observational studies is data-dredging; the inappropriate use of data mining to uncover misleading relationships in data (44). Researchers often set up studies with a large number of variables which can be related to a large number of outcomes, with one in 20 of the associations examined being statistically significant ($p < 0.05$) and thus acceptable for publication in medical journals (44). When a large number of associations can be looked at in a dataset where only a few real associations exist, the majority of findings will be false positives. These false positive findings are the true products of data dredging, resulting from simply looking at too many possible associations (44). In our cross-sectional study we demonstrated an association between JPC and health status (17). There is little to no transparency for the reader whether we tested this relationship on the basis of an *a priori* hypothesis – which we of course did – or that we tested multiple hypotheses and reported only on the positive associations. Since 2005, researchers are obliged to register their clinical trials (45), making it impossible for researchers to deviate from their *a priori* hypotheses and methods without disclosing this. At the moment, such registries are not mandatory for observational studies, leaving room for data dredging.

Pilot studies

In this thesis we assessed the feasibility of two different interventions by use of two different (pilot) study designs; a concurrent multiple baseline single case (MBSC) design and a randomized pilot trial. Both study designs have their own advantages and disadvantages in generating data on feasibility prior to a randomized clinical trial. A major advantage of the MBSC study is that it provides information on the health status of patients throughout the intervention, whereas the current tendency in randomized trials is still to provide only pre, post and follow-up measurements. Therefore, the information gathered in an MBSC study produces more information on the impact of the intervention during its course. Moreover, generally four or five people are enough to power this design, making it an attractive study design regarding time and costs.

However, an MBSC does not provide information on the logistics surrounding an intervention, for instance regarding recruitment rate or drop-out rate. Moreover, if an intervention is to be evaluated in a large scale randomized clinical trial, an MBSC study provides no data regarding the willingness of patients to undergo randomization. And finally, the high volume of evaluation moments in an MBSC study design can be emotionally tiresome for the participating volunteers (31). None of the disadvantages of the MBSC design apply for the randomized pilot study. But still, given the expenses of traditional trials and current budgetary constraints, researchers should explore different methods to efficiently test health care interventions or processes. In this case, the multiple baseline design might be an interesting alternative to the randomized pilot trial, especially if the researchers are particularly interested in testing the intervention.

12.3 IMPLICATIONS FOR FUTURE PRACTICE

Throughout the last couple of years there is heavy debate on the way we should look at health and the way we should manage disease (46-48). At a recent conference of the Health Council of the Netherlands, the following view on health was reached: “*the ability to adapt and to self-manage*” (49). This view on health corresponds well to core business of the health professional, namely to help individuals manage the consequences of their symptoms.

One of the most prominent consequences of OA is the loss in daily functioning. For health care providers to be able to target functional status, it is mandatory that functional status can be measured in a valid and reliable way. In the OA guideline of the Royal Dutch Society for Physical Therapy the HOOS/KOOS and Algofunctional Index are recommended to monitor and evaluate functioning (50). Considering that the Algofunctional Index measures both pain and functioning and that the WOMAC, HOOS and KOOS physical function scales do not have the ability to differentiate well between

pain and functioning (13;15;16;51), we therefore recommend the use of the LEFS to measure self-reported functioning in people with hip or knee OA (16).

Current clinical practice lacks a differentiated approach towards the management of individuals with OA (26). Having (joint pain) comorbidities can be an additional burden for people with OA (17;52-54). Our research findings indicate that current clinical practice manages patients with or without JPC similarly; with the exception of prescribing stronger pain killers (17;26). This approach seems unlikely to help people 1) reach the capacity to fulfil their potential and obligations, 2) manage their life, and 3) participate in social activities (48). To help individuals attain these goals a personalized, self-management program seems to be the intervention of choice (1;55). Unfortunately, scientific knowledge on the management of joint pain comorbidities in people with OA is lacking. Health professionals can therefore not rely on evidence-based practice. In these cases, health professional should rely on best-practice. Identifying patients who report JPC can be established fairly easily by use of straight-forward questions or a manikin (56). Consequently, results from our Delphi study provide a number of key elements that could aid health professionals obtain best practice (32), in which setting a functional diagnosis (57), monitoring outcome (58), and coming to individual and contextual relevant patient goals should be mandatory (59).

Preoperative functioning is strongly associated with postoperative functioning (19-24). It would therefore be of value to identify individuals with low preoperative functional status, pleading for a thorough functional screening as soon as an individual is scheduled for surgery. Several clinical prediction rules that allow health professionals to preoperatively identify individuals at risk for delayed functional recovery during the hospital-stay after total knee and hip replacement surgery are (becoming) available (37;60). However, initially 'fit' patients could worsen during the waiting list; potentially 25% according to our data (25). These individuals should be identified as soon as possible to help these persons cope with these changing circumstances (i.e. surgery). Unfortunately, in our review we were unable to identify predictors for this physical decline during the waiting list. Perhaps, in case of lengthy waiting lists it might be of value to monitor physical functioning more closely to detect changes in functional status (58); ideally by use of performance-based measures. Health professionals could then identify those cases where functional capacity deteriorates and try to prevent or even reverse this decline. Early preoperative exercise has demonstrated to be more effective than managing the consequences/complications in a later stadium in other types of diseases (61;62). Whether such an intervention would be opportune prior to major joint replacement remains to be seen, as additional research is necessary to establish data on the health gains in the frailest of elderly and to establish cost-effectiveness data (32;33;63). However, the lack of such data should not dissuade us from personalizing the currently - mainly one-size-fits-all orientated (64) - preoperative and postoperative trajectories. Current practice would be greatly improved when health

professionals know which people at risk for delayed recovery and can adjust their clinical practice accordingly.

12.4 RECOMMENDATIONS FOR FURTHER RESEARCH

We need more insight in the nature of comorbidities in OA. Throughout this thesis we refer to participants that have multiple painful joints in some studies as ‘individuals reporting JPC’ (17;18) and in other studies as ‘individuals diagnosed with GOA’ (30;31). Since, no (international) consensus is reached on a uniform definition for GOA (30) and no data is available which disease-mechanisms cause JPC in individuals with hip and knee OA (17;65), we are unable to draw conclusions on whether these samples overlap. Cross-sectional studies are needed to elucidate the nature of JPC as this could provide new starting points for treatment. Furthermore, intervention studies are needed on the management of people with OA and JPC as currently no efficacy data are available.

We believe that the effectiveness of a high intensity (33), home-based (63), preoperative therapeutic exercise on postoperative functional recovery and social participation should be evaluated in people at-risk for delayed functional recovery (37;60). Moreover, currently available clinical prediction rules (37;60) should be studied further to increase their predictive properties (perhaps by the implementation of psychosocial factors) and generalizability and to warrant their implementation in clinical practice (66).

Finally, international consensus should be reached on what aspects determine the quality of a therapeutic intervention. In this thesis we have shown that therapeutic quality appears to be of concern as none of the investigated studies met the majority of our predefined quality criteria (32). It is conceivable that the quality of a therapeutic intervention determines the measured effect in a trial and therefore an important factor for systematic reviews and meta-analyses. We believe a first step would be to determine what criteria are obligatory for therapeutic quality in non-pharmacological care. This consensus could be reached by use of an international Delphi study. Subsequently, authors should be stimulated to report these aspects of their intervention in their manuscripts. This could be done by scientific journals as well as the (inter)national health organizations, such as the World Confederation for Physical Therapy (WCPT). This way, the quality of rehabilitative interventions might be improved, but also the gap between science and clinical practice could be closed.

REFERENCE LIST

- (1) Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011;(9783):2115-26.
- (2) Hunter DJ, Hellio Le Graverand-Gastineau MP. How close are we to having structure-modifying drugs available? *Rheum Dis Clin North Am* 2008;(3):789-802.
- (3) Matthews GL, Hunter DJ. Emerging drugs for osteoarthritis. *Expert Opin Emerg Drugs* 2011;16(3):479-91.
- (4) Fransen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev*. 2008;(4):CD004376.
- (5) Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database Syst Rev*. 2009;(3):CD007912.
- (6) Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage*. 2008; 16(2):137-62.
- (7) Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK et al. OARSI recommendations for the management of hip and knee osteoarthritis Part III: changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage*. 2010;18(4):476-99.
- (8) Conaghan PG, Birrell F, Burke M, Cumming J, Dickson J, Dieppe P et al. osteoarthritis: national clinical guideline for care and management in adults. 2008.
- (9) Wylde V, Dieppe P, Hewlett S, Learmonth ID. Total knee replacement: is it really an effective procedure for all? *Knee* 2007;(6):417-23.
- (10) Hawker GA. Who, when, and why total joint replacement surgery? The patient's perspective. *Curr Opin Rheumatol*. 2006;(5):526-30.
- (11) Nilsson AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord*. 2003;4:10.
- (12) Bombardier C, Melfi CA, Paul J, Green R, Hawker G, Wright J et al. Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. *Med Care*. 1995;(4 Suppl):AS131-AS144.
- (13) Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JJ. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index. *Clin Rheumatol*. 2002;(6):466-71.
- (14) Terwee CB, Mokkink LB, Steultjens MP, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology (Oxford)*. 2006;(7):890-902.
- (15) Soderman P, Malchau H. Validity and reliability of Swedish WOMAC osteoarthritis index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). *Acta Orthop Scand*. 2000;(1):39-46.
- (16) Hoogboom TJ, de Bie RA, den Broeder AA, Van den Ende CH. The highly reliable, valid and responsive Dutch Lower Extremity Functional Scale is the preferred measure for

- functioning in hip or knee osteoarthritis: a validation study. *BMC Musculoskelet Disord*. 2012;13:117.
- (17) Hoogeboom TJ, den Broeder AA, Swierstra BA, de Bie RA, Van den Ende CH. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study. *Arthritis Care Res (Hoboken)* 2012; 64(1):54-8.
 - (18) Hoogeboom TJ, den Broeder AA, de Bie RA, Van den Ende CH. Longitudinal study of the impact of joint-pain comorbidities on quality of life and activity levels in people with knee osteoarthritis: Data from the Osteoarthritis Initiative. Submitted 2012.
 - (19) Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol*. 2002;29(6):1273-9.
 - (20) Kennedy LG, Newman JH, Ackroyd CE, Dieppe PA. When should we do knee replacements? *Knee* 2003;(2):161-6.
 - (21) Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am* 2004;(10):2179-86.
 - (22) Fitzgerald JD, Orav EJ, Lee TH, Marcantonio ER, Poss R, Goldman L et al. Patient quality of life during the 12 months following joint replacement surgery. *Arthritis Rheum* 2004;(1):100-9.
 - (23) Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI, Arenaza JC et al. Effect of patient characteristics on reported outcomes after total knee replacement. *Rheumatology (Oxford)* 2007;(1):112-9.
 - (24) Zeni JA, Jr., Snyder-Mackler L. Preoperative predictors of persistent impairments during stair ascent and descent after total knee arthroplasty. *JBJS Am* 2010;(5):1130-6.
 - (25) Hoogeboom TJ, Van den Ende CH, van der Sluis G, Elings J, Dronkers JJ, Aiken AB et al. The impact of waiting for total joint replacement on pain and functional status: a systematic review. *Osteoarthritis Cartilage* 2009; 17(11):1420-7.
 - (26) Hoogeboom TJ, Snijders GF, Cats HA, de Bie RA, Bierma-Zeinstra SM, van den Hoogen FH et al. Prevalence and predictors of health care use in patients with early hip or knee osteoarthritis: two-year follow-up data from the CHECK cohort. *Osteoarthritis Cartilage*. 2012 Jun;20(6):525-31.
 - (27) Black C, Clar C, Henderson R, MacEachern C, McNamee P, Quayyum Z et al. The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation. *Health Technol Assess* 2009; 13(52):1-148.
 - (28) Wandel S, Juni P, Tendal B, Nuesch E, Villiger PM, Welton NJ et al. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010;c4675.
 - (29) Stukstette M, Hoogeboom T, de RR, Koelmans P, Veerman E, den BA et al. A multidisciplinary and multidimensional intervention for patients with hand osteoarthritis. *Clin Rehabil* 2012; 26(2):99-110.
 - (30) Hoogeboom TJ, Stukstette MJ, de Bie RA, Cornelissen J, den Broeder AA, Van den Ende CH. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. *BMC Musculoskelet Disord* 2010; 11:142.

- (31) Hoogeboom TJ, Kwakkenbos L, Rietveld L, den Broeder AA, de Bie RA, Van den Ende CH. Piloting and preliminary evaluating a multi-disciplinary care program for individuals with generalized osteoarthritis on pain and self-efficacy: a concurrent multiple-baseline single case study. *BMJ Open*. 2012 Jul 19;2(4). pii: e001161.
- (32) Hoogeboom TJ, Oosting E, Vriezekolk JE, Veenhof C, Siemonsma PC, de Bie RA et al. Effectiveness of preoperative therapeutic exercise for major joint replacement on postoperative functional recovery: a meta-analysis and best practice evaluation. *PLoS One*. 2012;7(5):e38031.
- (33) Hoogeboom TJ, Dronkers JJ, Van den Ende CH, Oosting E, Van Meeteren NL. Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized pilot trial. *Clin Rehabil* 2010; 24(10):901-10.
- (34) Terwee CB, van der Slikke RM, van Lummel RC, Benink RJ, Meijers WG, de Vet HC. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. *J Clin Epidemiol* 2006;(7):724-31.
- (35) Stevens-Lapsley JE, Schenkman ML, Dayton MR. Comparison of self-reported knee injury and osteoarthritis outcome score to performance measures in patients after total knee arthroplasty. *PM R* 2011; 3(6):541-9.
- (36) Wright AA, Cook CE, Baxter GD, Garcia J, Abbott JH. Relationship between the Western Ontario and McMaster Universities Osteoarthritis Index Physical Function Subscale and physical performance measures in patients with hip osteoarthritis. *Arch Phys Med Rehabil* 2010; 91(10):1558-64.
- (37) Elings J, Chorus AM, de Gast A, Bausch-Goldbohm S, Van Meeteren NL. Preoperative prediction of clinical postoperative functional recovery in patients undergoing total hip replacement: The value of physical performance indices. Submitted 2012.
- (38) Howard JS, Mattacola CG, Howell DM, Lattermann C. Response shift theory: an application for health-related quality of life in rehabilitation research and practice. *J Allied Health* 2011; 40(1):31-8.
- (39) Robling M, Hood K. Response shift, responsiveness or recall bias? *Br J Gen Pract* 2002;(480):585.
- (40) Schwartz CE, Andresen EM, Nosek MA, Krahn GL. Response shift theory: important implications for measuring quality of life in people with disability. *Arch Phys Med Rehabil* 2007;(4):529-36.
- (41) Pasquina P, Tramer MR, Walder B. Prophylactic respiratory physiotherapy after cardiac surgery: systematic review. *BMJ* 2003; 327(7428):1379.
- (42) Herbert RD, Bo K. Analysis of quality of interventions in systematic reviews. *BMJ* 2005; 331(7515):507-9.
- (43) Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. *Ann Rheum Dis*. 2009; 68(9):1413-9.
- (44) Smith GD, Ebrahim S. Data dredging, bias, or confounding. *BMJ* 2002; 325(7378):1437-8.
- (45) Working Group on Best Practice for Clinical Trials Registers (BPG). 2012. 26-3-2012.
- (46) Jadad AR, O'Grady L. How should health be defined? *BMJ* 2008; 337:a2900.
- (47) What is health? The ability to adapt. *Lancet* 2009; 373(9666):781.

- (48) Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, Kromhout D et al. How should we define health? *BMJ* 2011; 343:d4163.
- (49) Health Counsel of the Netherlands. Invitational Conference 'Is health a state or an ability? Towards a dynamic concept of health'. 2010.
- (50) Peter WFH, Jansen MJ, Bloo H, Dekker-Bakker LMMCJ, Dilling RG, Hilberdink WKHA et al. *KNGF Guideline* for Physical Therapy in patients with Osteoarthritis of the hip and knee. Supplement to the Dutch Journal of Physical Therapy 2010; 120(1).
- (51) Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004;(4):335-41.
- (52) van Dijk GM, Veenhof C, Lankhorst GJ, Dekker J. Limitations in activities in patients with osteoarthritis of the hip or knee: the relationship with body functions, comorbidity and cognitive functioning. *Disabil Rehabil* 2009;(20):1685-91.
- (53) van Dijk GM, Veenhof C, Schellevis F, Hulsmans H, Bakker JP, Arwert H et al. Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. *BMC Musculoskelet Disord* 2008;95.
- (54) Reeuwijk KG, de RM, van Dijk GM, Veenhof C, Steultjens MP, Dekker J. Osteoarthritis of the hip or knee: which coexisting disorders are disabling? *Clin Rheumatol* 2010; 29(7):739-47.
- (55) Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. *Lancet* 2004; 364(9444):1523-37.
- (56) van den Hoven LH, Gorter KJ, Picavet HS. Measuring musculoskeletal pain by questionnaires: the manikin versus written questions. *Eur J Pain* 2010; 14(3):335-8.
- (57) Cieza A, Hlifikier R, Chatterji S, Kostanjsek N, Ustun BT, Stucki G. The International Classification of Functioning, Disability, and Health could be used to measure functioning. *J Clin Epidemiol* 2009; 62(9):899-911.
- (58) Glasziou P, Irwig L, Mant D. Monitoring in chronic disease: a rational approach. *BMJ* 2005; 330(7492):644-8.
- (59) Stucki G, Sangha O. Principles of rehabilitation. In: Klippel JH, Dieppe PA, editors. *Rheumatology*. 2nd ed. London, England: Mosby; 1998. 1-14.
- (60) Judge A, Javaid MK, Arden NK, Cushnaghan J, Reading I, Croft P et al. A clinical tool to identify patients who are most likely to receive long term improvement in physical function after total hip arthroplasty. *Arthritis Care Res (Hoboken)* 2012.
- (61) Hulzebos EH, Helders PJ, Favie NJ, de Bie RA, Brutel de la Riviere A, Van Meeteren NL. Preoperative intensive inspiratory muscle training to prevent postoperative pulmonary complications in high-risk patients undergoing CABG surgery: a randomized clinical trial. *JAMA* 2006;(15):1851-7.
- (62) Dronkers JJ, Lamberts H, Reutelingsperger IM, Naber RH, Dronkers-Landman CM, Veldman A et al. Preoperative therapeutic programme for elderly patients scheduled for elective abdominal oncological surgery: a randomized controlled pilot study. *Clin Rehabil* 2010; 24(7):614-22.
- (63) Oosting E, Jans MP, Dronkers JJ, Naber RH, Dronkers-Landman CM, Appelman-de Vries SM et al. Preoperative Home-Based Physical Therapy Versus Usual Care to Improve

Functional Health of Frail Older Adults Scheduled for Elective Total Hip Arthroplasty: A Pilot Randomized Controlled Trial. *Arch Phys Med Rehabil.* 2012;93(4):610-6.

- (64) Oosting E, Hooijboom TJ, Elings J, van der Sluis G, Van Meeteren NL. Inhoud en methodologische kwaliteit van ziekenhuisprotocollen fysiotherapie na totale kniearthroplastiek. *Nederlands Tijdschrift voor Fysiotherapie* 2009;119(6):186-92.
- (65) Suri P, Morgenroth DC, Kwok CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. *Arth Care Res (Hoboken).* 2010;62(12):1715-23.
- (66) McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA* 2000; 284(1):79-84.

Osteoarthritis is the most prevalent chronic joint disease. In the Netherlands, over half a million people suffer from this disease. Because osteoarthritis results in pain, stiffness, loss of function and reduced quality of life, it has a considerable impact on people's lives and on the society. No cure is yet available for people suffering from osteoarthritis. Having people to rely on therapies that aim to relieve their symptoms, keep them physically active, help them cope with their disease, and postpone their surgery. The effects of these therapies are often disappointing as they frequently result in relatively small health benefits. A growing body of evidence suggests osteoarthritis could be considered more like a heterogeneous disease regarding its pathology and clinical manifestation. For allied health care providers this would mean focussing even more on the clinical and functional diagnosis (i.e. what are the clinical manifestations, what are the person's capabilities, and what are the specific health needs) rather than focussing merely on the medical diagnosis (in this case, osteoarthritis).

In this thesis we first aimed to identify health needs of patients with osteoarthritis that might impact therapy results. Consequently, we assessed the feasibility and effectiveness of therapies - as often provided by allied health providers - that were specifically tailored to those health needs.

Health needs in people with osteoarthritis.

Often studied health needs in osteoarthritis are pain, stiffness and functional limitations. These health needs are typical for osteoarthritis and are often assessed using the Western Ontario and McMaster University Osteoarthritis index (WOMAC) or the Hip and Knee Osteoarthritis Outcomes Scores (HOOS and KOOS, respectively). In Chapter 2, we demonstrate that these scales should probably not be used to measure physical functioning as they lack discriminant validity (in other words, they cannot differentiate pain and physical functioning). An alternative to these questionnaires might be the Lower Extremity Functional Scale (LEFS), as it demonstrated high construct validity, good reliability, superior responsiveness and discriminant validity.

A patient group that we had particular interest in was people with osteoarthritis who report joint pain comorbidity (JPC). In chapters 3 and 4 we showed, using both cross-sectional and longitudinal data, that individuals reporting JPC experience a higher clinical burden than individuals who do not report JPC. Interestingly, we were unable to demonstrate relevant differences in health care consumption between the two groups, implying that the current health care system does not act accordingly.

Another group of individuals that we deemed at risk for compromised health status are those awaiting elective total hip or knee replacement. In chapter 5, by use of systematic review we demonstrated that the majority of people do not experience deterioration in pain or self-reported functional status while waiting for surgery. However approximately 25% of the patients did report clinically relevant deterioration during the waiting time. Perhaps

these individuals might benefit from a preoperative therapeutic exercise program to prepare them for surgery.

Finally, in Chapter 6, we looked at health needs in the context of the health care system. Ideally, individuals use health care modalities based on their discomfort level (i.e. disease-related factors) rather than based on predisposing (e.g. sex, age, ethnicity, etc) or enabling (e.g. insurance type, income, etc) factors. We found that health care utilization was low during the first two years of the disease. Only the use of supplements increased over time, even though their efficacy is questioned. It was remarkable though that predisposing and enabling factors contributed almost equally to health care use as disease-related factors. The latter implies that current health care practice does not take health needs into account.

Tailored care in people with osteoarthritis.

Before launching scientific research studies, we believe the studied intervention should first be thoroughly scrutinized. In chapter 7 we presented a systematic approach on how to develop an evidence-based therapy program by synthesizing information from expert opinions, patient perspectives, and scientific literature. According to this method, in Chapter 8, we developed a multi-disciplinary and multi-dimensional intervention for individuals with generalized OA (GOA). People were considered having GOA when they experienced complaints in three or more joint groups, and had at least two objective signs that indicate OA in at least two of those joints. In chapter 9 we evaluated the feasibility and potential effectiveness of a preliminary version of the treatment protocol by use of a concurrent, multiple baseline, single case study. Although no adverse events occurred and adherence rate was high, participant satisfaction about the therapy was moderate as well as the participants' estimate of effectiveness of the intervention. Moreover, we found no beneficial effects on pain or on self-efficacy. These data allowed us to improve the therapy program as studied in a randomized clinical trial (Chapter 8).

We hypothesized that it might be of value to optimize the preoperative functional status of individuals with high risk for disappointing outcomes after joint replacement surgery (Chapter 5). To investigate the validity of this hypothesis we systematically reviewed the available literature on the effectiveness of preoperative exercise, in Chapter 10. We found that none of the studied interventions fulfilled our criteria to be considered valid, leaving us unable to sound draw conclusions. Thus, it remains to be seen whether a high quality study would actually result in beneficial effects. In order to gather data needed to develop such a large, high-quality clinical trial, we pilot embarked on a pilot study. In Chapter 11, we describe a high intensity exercise program at an outpatient physical therapy department in individuals that were predisposed for delayed postoperative recovery. We found, that we were unable to include the population we envisioned prior to the study, as the eligible non-volunteers were both older and frailer. Non-participation was mainly cause by logistic issues; therefore not warranting a large randomized clinical trial in the same setting.

In Chapter 12, the general discussion, we describe our main findings and methodological reflections. We also discuss the implications of our work for future practice and we provide recommendations for future research. In short, we came to the following recommendations for future research:

- Insight in the nature of JPC in OA is needed.
- International consensus on a uniform definition for GOA should be reached;
- High quality studies on the management of people with GOA are needed;
- A therapeutically valid preoperative exercise programme in high-risk patients in the home situation should be evaluated; and
- International consensus on what aspects determine the quality of a therapeutic intervention should be reached.

Artrose is de meest voorkomende chronische gewrichtsaandoening. In Nederland lijdt meer dan een half miljoen mensen aan deze aandoening. Omdat artrose resulteert in pijn, stijfheid, verlies van functie en lagere kwaliteit van leven, heeft het een substantiële impact op het persoonlijke en maatschappelijke functioneren. Op dit moment bestaat er nog geen curatieve behandeling voor artrose. Om die reden zijn behandelopties gericht op verlichten van symptomen, behouden van fysieke activiteit, leren omgaan met de ziekte, en uitstellen van de operatie. De effecten van deze behandelingen zijn vaak teleurstellend, aangezien ze voor veel mensen resulteren in relatief kleine gezondheidsverbeteringen.

Verscheidene wetenschappelijke publicaties suggereren dat artrose beschouwd moet worden als een heterogene ziekte gezien de etiologie en klinische manifestatie. Voor paramedici betekent dit dat zij zich vooral moeten richten op de klinische en functionele gevolgen van artrose (ofwel, de klinische verschijnselen, capaciteiten/mogelijkheden en de individuele zorgvraag van de patiënt). In dit proefschrift hebben we in eerste instantie getracht de zorgvraag van patiënten in kaart te brengen om het mogelijk te maken de behandeling optimaal op af te stemmen. Vervolgens hebben we de haalbaarheid en effectiviteit van verschillende behandelprogramma's afgestemd op de specifieke zorgvraag getoetst.

Zorgvraag van mensen met artrose.

Pijn, stijfheid en functionele beperkingen zijn de meest onderzochte klachten bij artrose. Dit zijn typische artroseklachten en worden vaak geëvalueerd met de *Western Ontario and McMasters University Osteoarthritis index* (WOMAC) of met de *Hip en Knee Osteoarthritis Outcomes Scores* (HOOS en KOOS, respectievelijk). In hoofdstuk 2 tonen we aan dat deze schalen niet goed bruikbaar zijn om fysiek functioneren te meten omdat ze fysiek functioneren niet goed kunnen onderscheiden van pijn. Een alternatief voor deze vragenlijsten is mogelijk de *Lower Extremity Functional Scale* (LEFS), aangezien deze schaal - naast een hoge constructvaliditeit, een goede betrouwbaarheid, een superieure responsiviteit - pijn en functioneren beter onderscheidt.

Een groep personen waarin we geïnteresseerd waren, waren mensen met artrose en pijn in andere gewrichten (PAG). In hoofdstukken 3 en 4 lieten we zien, met behulp van transversale en longitudinale onderzoeksgegevens, dat mensen met PAG een grotere gezondheidslast ervaren dan mensen die geen PAG rapporteerden. Een interessante bevinding was dat we geen relevante verschillen vonden in zorggebruik tussen de twee groepen. Mogelijk betekent dit dat er in de huidige zorg geen rekening wordt gehouden met de aanwezigheid van PAG.

Een andere groep personen waarin we geïnteresseerd waren, waren mensen op de wachtlijst voor een electieve totale heup- of knieprothese. We veronderstelden dat deze groep mensen een groter risico had achteruit te gaan in hun gezondheid tijdens de wachtlijst. Middels een systematische review, toonden we in hoofdstuk 5 aan dat de meerderheid van de mensen

geen verslechtering in pijn of zelf-gerapporteerde functionele status ervoeren tijdens het wachten op de operatie. Echter, 25% van de patiënten rapporteerde klinisch relevante achteruitgang in functioneren tijdens de wachttijd. Het is waarschijnlijk dat juist deze personen profiteren van een pre-operatief, therapeutisch oefenprogramma voorafgaand aan de operatie om hen voor te bereiden op de operatie en sneller te laten herstellen na de operatie.

In hoofdstuk 6 is gekeken naar zorggebruik bij artrose. Idealiter maken mensen gebruik van de gezondheidszorg op basis van ziekte-gerelateerde factoren (zoals, pijn, stijfheid en functieverlies) in plaats van persoonsrelateerde factoren (zoals geslacht, leeftijd, etniciteit, type verzekering, inkomen, etc). Uit ons cohortonderzoek bleek dat het zorggebruik van artrosepatiënten laag was gedurende de eerste twee jaar van de ziekte. Alleen het gebruik van supplementen nam toe over de tijd. Opvallend was dat persoonsrelateerde factoren in gelijke mate bijdroegen aan zorggebruik als ziekte-gerelateerde factoren. Dit impliceert dat er ongelijkheid bestaat in de toegankelijkheid van zorg voor mensen met artrose.

Zorg op maat voor mensen met artrose.

In hoofdstuk 7 presenteerden we een systematische aanpak voor de ontwikkeling van een evidence-based therapie door informatie van deskundigen, patiënten perspectieven en wetenschappelijke literatuur te combineren. In hoofdstuk 8 ontwikkelden we een multidisciplinair en multi-dimensionaal behandelprotocol voor mensen met een gegeneraliseerde artrose op basis van deze systematische aanpak. Gegeneraliseerde artrose werd gedefinieerd als “klachten in drie of meer gewrichtsgroepen en ten minste twee objectieve artroseverschijnselen in tenminste twee van deze gewrichten”. In hoofdstuk 9 evalueerden we de haalbaarheid en de potentiële effectiviteit van een eerste versie van dit behandelprotocol met behulp van een multiple baseline, single case studie. Hoewel er geen bijwerkingen optraden en de naleving van het behandelprotocol hoog was, waren de deelnemers slechts matig tevreden over het behandelprogramma en schatten zij de effectiviteit van de interventie laag in. Bovendien vonden we geen gunstig effect op pijn of op self-efficacy, onze primaire uitkomsten. Op basis van de gegevens van het onderzoek is het behandelprotocol bijgesteld.

Zoals eerder beschreven (hoofdstuk 5), is het mogelijk zinvol om middels therapeutische training individuen met een hoog risico op teleurstellende resultaten na een gewrichtsvervangende operatie voor te bereiden op een operatie. Deze hypothese toetsten we middels een systematische literatuurstudie. Hiervoor beoordeelden we, in hoofdstuk 10, de beschikbare literatuur over de effectiviteit van preoperatieve training. We vonden dat geen van de onderzochte interventies voldeed aan onze voorafgestelde kwaliteitscriteria. Hierdoor waren we niet in staat conclusies te trekken. Een gerandomiseerde studie is nodig om te bepalen of een dergelijke hoog kwalitatieve interventie leidt tot positieve effecten. Om de gegevens te genereren die nodig zijn voor een dergelijke studie, startten we met een

pilot-studie zoals beschreven in hoofdstuk 11. In dit hoofdstuk onderzochten we een intensief oefenprogramma op een poliklinische fysiotherapieafdeling voor fragiele oudere personen, die mogelijk een hoog risico op vertraagd herstel na de operatie hadden. We vonden dat de meest fragiele ouderen niet wilden deelnemen aan ons onderzoek, met name vanwege logistieke overwegingen. Daarom kunnen wij een grote gerandomiseerde klinische studie in dezelfde instelling niet rechtvaardigen.

In hoofdstuk 12, de algemene discussie, beschrijven we onze belangrijkste bevindingen en methodologische reflecties. Verder bespreken we de implicaties van ons werk voor de toekomstige praktijk en doen we aanbevelingen voor toekomstig onderzoek. We kwamen tot de volgende aanbevelingen voor toekomstig onderzoek. Allereerst is er meer inzicht in de aard van PAG bij artrose nodig, opdat behandelprogramma's nog beter afgestemd op de deelnemers afgestemd kunnen worden. Verder dient er internationale consensus te worden bereikt over een uniforme definitie voor generaliseerde artrose, zodat er meer systematisch en reproduceerbaar (effectiviteits)onderzoek kan worden verricht in dit gebied. Tevens zijn er hoog kwalitatieve studies naar het effect van therapeutische preoperatieve trainen van hoog-risico patiënten aan huis nodig, daar deze patiëntengroep niet naar het ziekenhuis kan/wil komen. Tot slot is er een grote behoefte aan internationale consensus over welke aspecten de kwaliteit van een therapeutische interventie karakteriseren om de ontwikkeling en evaluatie van hoogwaardige kwalitatieve therapeutische oefenprogramma's wereldwijd te bevorderen.

Beste lezer, mocht u bij dit hoofdstuk aangekomen zijn na het lezen van alle voorgaande hoofdstukken; chapeau. Als auteur wens ik u jolijt en vreugd bij het lezen van dit laatste caput. Echter, mocht u direct doorgebladerd zijn naar dit dankwoord, zonder kennis te nemen van de verdere inhoud van dit proefschrift, dan hoop ik dat u zich realiseert dat mijn wetenschappelijke hart weent. Alle eerdere hoofdstukken van deze thesis belichamen jaren van arbeid, frustraties en wanbof, terwijl het huidige hoofdstuk vrij snel geschreven was. Ik hoefde hiervoor geen literatuurstudies te ondernemen, geen onderzoeksprojecten op te zetten en ook geen peer-review proces te doorstaan. Neen, ik hoefde voor dit hoofdstuk enkel de leuke herinneringen op te halen en terug te denken aan alle fijne mensen met wie ik heb samengewerkt in de afgelopen vier jaar.

Laat ik bij het begin beginnen, namelijk bij mijn promotieteam, bestaande uit *Els van den Ende*, *Alfons den Broeder* en *Rob de Bie*. Van deze drie personen heb ik veruit het meeste contact gehad met Els, mijn eerste copromotor. In eerste instantie lijkt de combinatie Els & Thomas misschien niet optimaal. Immers, hoe kan een ‘strikte’, ‘streng’, en doorgewinterde onderzoekster in vredesnaam samenwerken met een naïeve, onervaren en impulsieve jonge hond? Voordat ik überhaupt begon ging het al bijna fout, omdat ik te laat en veel te informeel aan Els vroeg of ik nog mocht solliciteren. Echter, het is ons toch gelukt. We zijn productief geweest, hebben een aantal mooie artikelen gepubliceerd en we hebben samen een hoop lol gehad (alhoewel, jij met name ten koste van mij). Els, ik heb zoveel van je geleerd. Jij hebt mij opgeleid tot een complete en kritische onderzoeker. Bedankt! Mijn tweede copromotor, Alfons, wil ik met name bedanken voor zijn interessante en fijne feedback. Toen ik Alfons vroeg het manuscript over de validatiestudie van de LEFS (hoofdstuk 2) te bekijken, kreeg ik naast zinvol commentaar, ook de top 5 beste klinimetriegrappen toegestuurd (zoals ‘Zegt een klinimetrist tegen twee boeven: “Is die maat van jou wel betrouwbaar?”’). Alfons, bedankt voor de leuke en uitdagende discussies en je bijdrage aan mijn plezierige tijd in de Sint Maartenskliniek. En natuurlijk wil ik mijn promotor, Rob, bedanken. Rob, de eerste keer ik jou ontmoette (op het congres Onderzoek in Beweging in Maastricht) voelde ik een soort klik. Is dit wat ze noemen ‘liefde op het eerste gezicht’? Blijkbaar dacht jij daar hetzelfde over, immers je gaf aan mij als PhD-student te willen begeleiden. Bovendien gaf je mij zes maanden onderdak aan de universiteit Maastricht toen ik na 3 jaar vanwege de flexwet tijdelijk weg moest bij de Sint Maartenskliniek. Rob, bedankt voor je begeleiding tijdens mijn promotie, je eindeloze grappen over raften en je supersnelle reacties op mijn emails. Els, Alfons en Rob, ik hoop in de toekomst nog vaak met jullie samen te werken!

Vervolgens wil ik graag de leden van de beoordelingscommissie bedanken. Prof. dr. *Sjef van der Linden*, prof. dr. *Sita Bierma-Zeinstra*, prof. dr. *Andre Knottnerus*, dr. *Nico van Meeteren* en prof. dr. *Martin Prins* bedankt voor het lezen en beoordelen van mijn proefschrift.

Verder wil ik al mijn collega’s bedanken in de Sint Maartenskliniek. Allereerst de Reumaresearchers. Bedankt voor jullie kritische commentaren en gezelligheid. *Agnes*, mijn

tijd met jou samen op de kamer zie ik als één van de leukste, creatiefste en productiefste periodes van mijn promotie. Het feest dat je had geregeld bij mijn afscheid was echt fantastisch. Je bent een ware topcollega en ik ben vereerd dat je mijn paranimf wilde zijn. *Gijs, Nienke, Mirelle, Linda en Joke*, allen bedankt voor jullie bijdrage als coauteurs aan verscheidene artikelen in dit proefschrift. *Dicky*, bedankt voor al je hulp in de logistiek van mijn onderzoeken. Zonder jullie had dit proefschrift nooit zijn huidige vorm bereikt. Ten tweede mijn onderzoekscollega's van RD&E; een gezellige groep mensen, die het werken op de Sint Maartenskliniek tot een (letterlijk) feest hebben verheven: *Amber, Astrid, Cheriël, Judith, Marianne, Niki, en Noël*. Verder, iedereen van RD&E (ondermeer *Bart & Jacques*) ontzettend bedankt voor jullie gastvrijheid! Met één uitzondering daar gelaten, *Koen Koenraadt*, wiens gefluit mij danig heeft belemmerd in het realiseren van dit proefschrift ;-). Ten derde, *Werner, Elke en Jan*, daar onze eindeloze gesprekken over fietsen, voetbal en Dries voor de broodnodige afleiding van mijn proefschrift zorgde. Verder wil ik dr. *Bart Swierstra* bedanken voor zijn hulp in de patiënteninclusie voor een aantal van mijn studies. En als laatste wil ik alle medewerkers van de Ambulant Reumacentrum (Nijmegen en Woerden) bedanken waarbij speciale dank voor de volgende personen: *Claudia, Eliane, Ineke, Janneke de B, Janneke S, Jeannette, Léon, Sylvia, en Yvette*. Een extra speciale dank aan *Clarinda* (ik heb ontzettend genoten van onze samenwerking aan de MEGA studie en onze tripjes door het land, hopelijk houden we contact!), *Leonie* (bedankt voor al je inzet voor mijn projecten in Woerden) en *Frank van den Hoogen* (zonder jouw passie en inzet zou Reumaresearch nooit zo een mooie onderzoeksgroep zijn geweest).

Natuurlijk zijn er ook een aantal mensen buiten de Maartenskliniek die ik graag wil bedanken voor hun (in)directe bijdrage aan dit proefschrift. Mijn collega's van de Bunnik groep - *Jordi Elings, Ellen Oosting en Geert van der Sluis* - bedankt voor alle gezellige, leerzame en productieve bijeenkomsten. Mijn oud-studiegenoten; speciaal *Pieter Schipper* (mijn nachtpromotor) en *Peter van Essen* (de beste databeukelaar ooit). Mijn familie en vrienden; speciaal *Michiel Ars* (mijn tweede paranimf en goede vriend; wat moet Zeeland zonder jou?), *Stefan Hopmans* en *Lard Hermse* (PES in de Muts!) en *Vera van Schagen* (vanwege je vriendschap en je vertaalhulp in hoofdstuk 2). Verder wil ik graag *Jaap Dronkers, Jennifer Stevens-Lapsley, Paul Stratford, Cindy Veenhof, Janet Wesseling* (en de gehele CHECK onderzoeksgroep) en *Conny de Zwart* bedanken voor al hun bijdragen aan de totstandkoming dit proefschrift. Tot slot wil ik *Nico van Meeteren* bedanken wiens visie op het vak fysiotherapie mij altijd inspireert; ik kijk dan ook uit naar onze toekomstige samenwerking.

Ook wil ik mijn ouders, *Hans en Truus*, ontzettend bedanken voor hun steun gedurende mijn hele opleiding en promotie. Dankzij jullie treed ik ooit nog eens in Hans zijn voetsporen. De opmerking van mijn vader toen ik hem trots een kopij van mijn proefschrift overhandigde: "Op de HTS schrijven studenten dikkere verslagen in 5 maanden tijd", motiveert mij om tijdens mijn post-doc positie me nog beter in te zetten... Bedankt pa ;-).

Tenslotte wil ik mijn vriendin *Sanne* bedanken voor al haar hulp, steun en liefde voor, tijdens en na mijn promotietraject. Ik hou van jou! Niet alleen heb jij deze promotieplek voor mij gevonden. Jij was er ook wanneer ik weer eens gestresst thuis kwam na een frustrerende PhD-dag (lees: na een afwijzing door een wetenschappelijk tijdschrift, na het vinden van niet-significante resultaten, na een dag dat je niet meer kunt herinneren wat je hebt gedaan, of na het contempleren over levensvragen als “Wie ben ik” en “Waar ga ik heen”). En ook tijdens de (vluchtige) hoogtepunten van mijn promotietraject (lees: na een acceptatie van een artikel, na de realisatie dat ik bijdroeg aan verwerving van wetenschappelijke kennis betreffende het menselijk functioneren in het belang van de maatschappij, of na een dag werken dat ik me wel kon herinneren wat ik had gedaan). Jouw liefde, kusjes en statistiek kennis hebben een grote bijdrage geleverd aan de realisatie van dit proefschrift. Ik hoop nog heel lang van al je kwaliteiten te mogen genieten; binnenkort als mijn vrouw.

Thomas J Hoogeboom was born on February 14, 1982 in Goes, the Netherlands. He completed his secondary school at the St. Willibrordcollege in Goes, and studied Physical Therapy at the Avans Hogeschool in Breda. After graduating in 2004 (BHS), Thomas resumed his education by studying Human Movement Sciences at the University Medical Centre St Radboud in Nijmegen. Throughout his education Thomas worked as a physical therapist in several primary care practices. Additionally he worked as a consultant for the electronic patient file Abakus® and as a researcher for University Medical Centre Utrecht and health insurance company AGIS. In February 2008 he graduated as a human movement scientist (MSc). Shortly after that (May 2008), he embarked on his PhD-trajectory at the Department for Rheumatology at the Sint Maartenskliniek in Nijmegen, in cooperation with the Department of Epidemiology at the Maastricht University Medical Centre+ (where he physically worked for 6 months (May - October 2011)). While working on his PhD, Thomas also worked as a guest lecturer at three Physical Therapy schools and a Master of Science program. In May 2012 he signed a contract with the Centre for Care Technology Research (CCTR) at the Maastricht University Medical Centre+ for a postdoc position. At that same time, he moved to the United States to work as a researcher at the University of Colorado in Aurora, Colorado. His current and future projects will focus on therapeutic validity in physical therapy interventions and care before, during, and after total joint replacement.

International scientific publications:

Hoogeboom TJ, de Bie RA, den Broeder AA, van den Ende CHM. The Dutch Lower Extremity Functional Scale was highly reliable, valid and responsive in individuals with hip/knee osteoarthritis: a validation study. BMC Musculoskeletal Disorders. 2012 Jul 2;13:117.

Hoogeboom TJ, den Broeder AA, de Bie RA, van den Ende CHM. Longitudinal study of the impact of joint-pain comorbidities on quality of life and activity levels in people with knee osteoarthritis: Data from the Osteoarthritis Initiative. Rheumatology (Oxford), 2012.

Hoogeboom TJ, Kwakkenbos L, Rietveld L, den Broeder AA, de Bie RA, van den Ende CHM. Feasibility and potential effectiveness of a non-pharmacological multidisciplinary care programme for persons with generalised osteoarthritis: a randomised multiple-baseline single-case study. BMJ Open. 2012 Jul 19;2(4).

Hoogeboom TJ, Oosting E, Vriezekolk JE, Veenhof C, Siemonsma PC, de Bie RA, van den Ende CH, van Meeteren NL. Therapeutic validity and effectiveness of preoperative exercise on functional recovery after joint replacement: a systematic review and meta-analysis. PLoS One. 2012;7(5):e38031.

Hoogeboom TJ, Snijders GF, Cats HA, de Bie RA, Bierma-Zeinstra SM, van den Hoogen FH, van Riel PL, Emans PJ, Wesseling J, den Broeder AA, van den Ende CH. Prevalence and predictors of health care use in patients with early hip or knee osteoarthritis: two-year follow-up data from the CHECK cohort. Osteoarthritis Cartilage. 2012;20(6):525-31.

Cuperus N, **Hoogeboom TJ**, Neijland Y, van den Ende CH, Keijsers NL. Are people with rheumatoid arthritis who undertake activity pacing at risk of being too physically inactive? Clin Rehabil. 2012 Feb 10.

Stukstette MP, **Hoogeboom TJ**, de Ruiter R, Koelmans P, Veerman E, den Broeder AA, Cats HA, Bijlsma JW, Dekker J, van den Ende CH. A multidisciplinary and multidimensional intervention for patients with hand osteoarthritis. Clin Rehabil. 2012;26(2):99-110.

Hoogeboom TJ, den Broeder AA, Swierstra BA, de Bie RA, van den Ende CH. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study. Arthritis Care Res (Hoboken). 2012;64(1):54-8.

Hoogeboom TJ, Stukstette MJ, de Bie RA, Cornelissen J, den Broeder AA, van den Ende CH. Non-pharmacological care for patients with generalized osteoarthritis: design of a randomized clinical trial. BMC Musculoskelet Disord. 2010;11:142.

Hoogeboom TJ, Dronkers JJ, van den Ende CH, Oosting E, van Meeteren NL. Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized pilot trial. Clin Rehabil. 2010;24(10):901-10.

Hoogeboom TJ, van den Ende CH, van der Sluis G, Elings J, Dronkers JJ, Aiken AB, van Meeteren NL. The impact of waiting for total joint replacement on pain and functional status: a systematic review. Osteoarthritis Cartilage. 2009;17(11):1420-7.

National scientific publications:

Hoogeboom TJ, Oosting E, Vriezekolk JE, Veenhof C, Siemonsma PC, de Bie RA, van den Ende CH, van Meeteren NL. Therapeutische validiteit en effectiviteit van preoperatieve therapeutische training voorafgaand aan een totale gewrichtsvervanging: een systematische review en meta-analyse. Nederlands Tijdschrift voor Fysiotherapie. 2012.

Oosting E, **Hoogeboom TJ**, Elings J, van der Sluis G, van Meeteren NL. Inhoud en methodologische kwaliteit van ziekenhuisprotocollen fysiotherapie na totale knieartroplastiek. Nederlands Tijdschrift voor Fysiotherapie. 2009;119(6):186-192.

Hoogeboom TJ, Elings J, van der Sluis G, Oosting E, van Meeteren NLU. Fysiotherapie in de ketenzorg voor patiënten die een totale heup- of knieartroplastiek ondergaan. Stimulus. 2008;27(3):223-237.

van der Sluis G, Elings J, **Hoogeboom TJ**, D'hondt NE, van Meeteren NLU. Fysiotherapie bij totale heup- of knieartroplastiek: ook een pré? Fysiopraxis. 2007;16(2):24-27.